

## Swope, Sheridan

**From:** Swope, Sheridan  
**Sent:** Saturday, August 19, 2006 4:30 PM  
**To:** Hill, Dwayne (ASRC)  
**Subject:** FW: 10/734,719

**Importance:** High

Dwayne, Unfortunately, this search was not done correctly.  
SID 9, 12, and 14 were done against the AA databases, not the NT databases.  
Also, I need some more hits from certain databases.  
So, would you put the following in as soon as you can?

For 10/734,719, pls search & interference search the following:  
SID 8 search only against the .rni, .rnpm, and .rge NT databases--and collect 150 hits.

SID 9 against all NT databases  
- for databases .rni, .rnpm, and .rge, collect 150 hits.

SID 12 against all NT databases  
SID 14 against all NT databases

Pls let me know when you think it will be done.  
THanks,  
SHeridan Swope  
ps, let me know if you have any questions.

*Handwritten notes:*  
rge ✓  
rni ✓  
rnpm ✓  
EST ✓  
rst ✓  
rnpm ✓  
rnpn ✓

\*\*\*\*\*8

-----Original Message-----

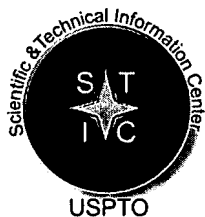
**From:** Swope, Sheridan  
**Sent:** Tuesday, July 18, 2006 1:10 PM  
**To:** STIC-Biotech/ChemLib  
**Subject:** 10/734,719

For 10/734,719, pls search & interference search the following:

SID 8 against the NT databases  
SID 9 against the NT databases  
SID 11 against the NT databases  
SID 12 against the NT databases  
SID 13 against the NT databases  
SID 14 against the NT databases

*Handwritten note:*  
Not 8/28/06

Sheridan Swope, Ph.D.  
Primary Patent Examiner  
AU 1656/Recombinant Enzymes  
571-272-0943 (voice)



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 195857**

**TO: Sheridan Swope**  
**Location: REM-2D19&3C70**  
**Art Unit: 1656**  
**6/23/06**  
**Case Serial Number: 10/734719**

**From: Dwayne Hill**  
**Location: Biotech-Chem Library**  
**REM-1A41**  
**Phone: (571)272-8972**

**Dwayne.Hill@uspto.gov**

### **Search Notes**

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Dwayne Hill  
STIC Biotech/Chem Library  
(571)272-1972



STIC-Biotech/ChemLib

195837

From: Swope, Sheridan  
Sent: Tuesday, July 18, 2006 1:10 PM  
To: STIC-Biotech/ChemLib  
Subject: 10/734,719

For 10/734,719, pls search & interference search the following:

SID 8 against the NT databases  
SID 9 against the NT databases  
SID 11 against the NT databases  
SID 12 against the NT databases  
SID 13 against the NT databases  
SID 14 against the NT databases

Sheridan Swope, Ph.D.  
Primary Patent Examiner  
AU 1656/Recombinant Enzymes  
571-272-0943 (voice)  
E02D19 Remsen Bld (Office)  
E03C70 Remsen Bld (Mailbox)  
Helping applicants get good patents.

CRFE

NO HIT  
8/8/06

79280

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 196857**

**TO: Sheridan Swope**  
**Location: REM-2D19&3C70**  
**Art Unit: 1656**  
**Thursday, July 20, 2006**  
**Case Serial Number: 10/734719**

**From: Dwayne Hill**  
**Location: Biotech-Chem Library**  
**REM-1A64**  
**Phone: (571) 272-8972**

**Dwayne.hill@uspto.gov**

### **Search Notes**

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Dwayne Hill  
Technical Information Specialist  
STIC Biotech/Chem Library  
(571) 272-8972

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: August 21, 2006, 14:42:06 ; Search time 4772.5 Seconds  
(without alignments)  
5848.727 Million cell updates/sec

Title: US-10-734-719-12  
Perfect score: 1563  
Sequence: 1 MKKVIIAGNGPSLKEIDYSR.....KLIKDLLRLPSDIKHVFGK 291

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues  
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q-/abs/ABSSWEB spool/US10734719/runat\_21082006\_141810\_12316/app\_query.fasta\_1  
-DB=GenEmbl -OPWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05p  
-NO MAP -NEG SCORES=0 -CGN 1.1 8663 @runat\_21082006\_141810\_12316 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb env.\*  
2: gb pat.\*  
3: gb ph.\*  
4: gb pl.\*  
5: gb pr.\*  
6: gb ro.\*  
7: gb sts.\*  
8: gb sy.\*  
9: gb un.\*  
10: gb vi.\*  
11: gb ov.\*  
12: gb htg.\*  
13: gb in.\*  
14: gb on.\*  
15: gb ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	1563	100.0	873	2 AX934434 Sequence
2	1563	100.0	873	2 AR481786 Sequence
3	1563	100.0	873	2 AR527385 Sequence

4	1563	100.0	873	2 AR609664	AR609664	Sequence
5	1563	100.0	873	2 AR689942	AR689942	Sequence
6	1563	100.0	873	2 AR691838	AR691838	Sequence
7	1563	100.0	11474	15 AF215659	AF215659	Campyloba
8	1557	99.6	876	2 AX934431	AX934431	Sequence
9	1557	99.6	876	2 BD249794	BD249794	Campyloba
10	1557	99.6	876	2 AR271703	AR271703	Sequence
11	1557	99.6	876	2 AR481785	AR481785	Sequence
12	1557	99.6	876	2 AR527384	AR527384	Sequence
13	1557	99.6	876	2 AR609663	AR609663	Sequence
14	1557	99.6	876	2 AR689941	AR689941	Sequence
15	1557	99.6	876	2 AR691837	AR691837	Sequence
16	1557	99.6	4749	15 AF305571	AF305571	Campyloba
17	1557	99.6	6047	15 AY661458	AY661458	Campyloba
18	1557	99.6	24425	15 AY422197	AY422197	Campyloba
19	1557	99.6	24437	15 AF167344	AF167344	Campyloba
20	1553	99.4	873	2 AX934436	AX934436	Sequence
21	1553	99.4	873	2 AR481787	AR481787	Sequence
22	1553	99.4	873	2 AR527386	AR527386	Sequence
23	1553	99.4	873	2 AR609665	AR609665	Sequence
24	1553	99.4	873	2 AR689943	AR689943	Sequence
25	1553	99.4	873	2 AR691839	AR691839	Sequence
26	1553	99.4	12390	15 AF401528	AF401528	Campyloba
27	1534	98.1	876	2 AX934427	AX934427	Sequence
28	1534	98.1	876	2 BD249792	BD249792	Campyloba
29	1534	98.1	876	2 AR271701	AR271701	Sequence
30	1534	98.1	876	2 AR481783	AR481783	Sequence
31	1534	98.1	876	2 AR527382	AR527382	Sequence
32	1534	98.1	876	2 AR609661	AR609661	Sequence
33	1534	98.1	876	2 AR689939	AR689939	Sequence
34	1534	98.1	876	2 AR691835	AR691835	Sequence
35	1531	98.0	11442	15 AF400048	AF400048	Campyloba
36	1524	97.5	876	2 AX934425	AX934425	Sequence
37	1524	97.5	876	2 BD249791	BD249791	Campyloba
38	1524	97.5	876	2 AR271700	AR271700	Sequence
39	1524	97.5	876	2 AR481782	AR481782	Sequence
40	1524	97.5	876	2 AR527381	AR527381	Sequence
41	1524	97.5	876	2 AR609660	AR609660	Sequence
42	1524	97.5	876	2 AR689938	AR689938	Sequence
43	1524	97.5	876	2 AR691834	AR691834	Sequence
44	1524	97.5	876	15 AF216647	AF216647	Campyloba
45	1524	97.5	6047	15 AY644679	AY644679	Campyloba

ALIGNMENTS

RESULT 1	AX934434	Sequence 11 from Patent WO02074942.	873 bp	DNA	linear	PAT 05-JAN-2004
LOCUS	AX934434	Sequence 11 from Patent WO02074942.	873 bp	DNA	linear	PAT 05-JAN-2004
DEFINITION	AX934434	Sequence 11 from Patent WO02074942.	873 bp	DNA	linear	PAT 05-JAN-2004
ACCESSION	AX934434	Sequence 11 from Patent WO02074942.	873 bp	DNA	linear	PAT 05-JAN-2004
VERSION	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
KEYWORDS	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
SOURCE	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
ORGANISM	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
REFERENCE	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
AUTHORS	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
TITLE	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
JOURNAL	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
FEATURES	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
source	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004
CDS	AX934434.1	GI:40641680	873 bp	DNA	linear	PAT 05-JAN-2004

```
/transl_table=11
/protein_id="CAF04170.1"
/db_xref="GI:40641681"
/translation="MKKVIILAGNPSLKEIDYSLRPNDFVPRCQNFEBDKYILGKK
CKAVYTFPFQFEQYTLKHLITQOEYETELIMCSNYQAHLENENFVKTFYDYPDA
HLGVDFFKQLKEFNAYKFHEIYFNQRTISGVYMCVAIAIYALGYKEIYILYDFYQNGS
SYAFTKQENLKLAPDFKNDKSHVIGHSKNTDIKALEFLEKTYIKLYCLCPNLLA
NFIELAPNLNSFIIOEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENVYKLIKD
LLRLPSDIKHFKGK"

ORIGIN
Alignment Scores:
Pred. No.: 3,09e-191 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AX934434 (1-873)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLysLeuLysAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGCAATGACCAAGTTTAAAAAGAAATTGATTATCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAAAAATGCAAGACAGTGTTTTACACCCCTGGTTCCTTTCTTGAGCAATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATAGAGCCGAACTAATATGTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAAAACTTTTACGATTTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGluLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAACCCTACAAAATAAAACTATATGCTTATGCTCTAACAGCTTTTGTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACCTAGCGCCCAATTTTAAATTTTCAATTTTATCATACAGAAAAATTAACCTACCT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACCTACCTCTTAGTGAGGCTTATGAGAAATTTTCAAAAAATATTAATTTT 780

RESULT 2
AR481786 LOCUS
DEFINITION Sequence 11 from patent US 6699705.
ACCESSION AR481786
VERSION AR481786.1 GI:47243421
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert, M. and Wakarchuk, W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6699705-A 11 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..873
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3,09e-191 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AR481786 (1-873)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLysAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGCAATGACCAAGTTTAAAAAGAAATTGATTATCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAAAAATGCAAGACAGTGTTTTACACCCCTGGTTCCTTTCTTGAGCAATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATAGAGCCGAACTAATATGTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAAAACTTTTACGATTTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAGCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAACCCTACAAAATAAAACTATATGCTTATGCTCTAACAGCTTTTGTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACCTAGCGCCCAATTTTAAATTTTCAATTTTATCATACAGAAAAATTAACCTACCT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACCTACCTCTTAGTGAGGCTTATGAGAAATTTTCAAAAAATATTAATTTT 780
```

```

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCCTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAAACATATATGCTTAGTCCTAACAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAAGCAATATTTCAAGGAAAA 873

RESULT 3
AR527385
LOCUS AR527385 873 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 11 from patent US 6723545.
ACCESSION AR527385
VERSION AR527385.1 GI:53914406
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 11 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..873
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 3.09e-191 Length: 873
Pred. No.: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AR527385 (1-873)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGAATGATTTAGATGATCAATATTTTATTGTAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGACAGTGTTTACACCCCTGGTTCTCTTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80

```

```

Db 181 ACTTTAAACATTTAATCCAAATATCAAGAAATATGAGACCGAACTAAATATATGTTCTTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTTCAGATTATTTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAAGAAATTTTCAATCAAGAAATTAACCTCAGGGGCTATATATGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCATTATGCTTTTGATACCAACAGAAAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTCAAAAAATAAACTATATGCTTAGTCCTTAACAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAAGCAATATTTCAAGGAAAA 873

RESULT 4
AR609664
LOCUS AR609664 873 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 11 from patent US 6825019.
ACCESSION AR609664
VERSION AR609664.1 GI:56664964
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 11 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..873
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 3.09e-191 Length: 873
Pred. No.: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0

```

Best Local Similarity: 100.0%	Mismatches: 0
Query Match: 100.0%	Indels: 0
DB: 2	Gaps: 0
US-10-734-719-12 (1-291) x AR609664 (1-873)	
Qy	1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db	1 ATGAAAGAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60
Qy	21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db	61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120
Qy	41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db	121 CTTGGTAAAAATGCAAGCAGTGTTTTACCCCTGGTTCTTCTTTGAGCAATACTAC 180
Qy	61 ThrLeuLysHisLeuLysGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db	181 ACTTTAAACACATTTAATCCAAATCAAGAATATGAGACCGAATTAATGTGTCTTAAT 240
Qy	81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100
Db	241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAAACTTTTATACGATTATTTTCC 300
Qy	101 AspAlaHisLeuLysGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 120
Db	301 GATGCTCATTTGGGATATGATTTTAAAACTTTTGTAAAACTTTTATACGATTATTTTCC 360
Qy	121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db	361 TTTTACCAAGATTTTAAATCAAGAAATACCTCAGGGTCTATATGTGTGCGAGTAGCC 420
Qy	141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyLysPheTyrGlnAsnGlySer 160
Db	421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAAATGGGTCA 480
Qy	161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db	481 TCTTATGCTTTTGATACCAACAGAAAACTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy	181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db	541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy	201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db	601 GAAAAAACTTACAAAATAAAACTATATTGCTTATGCTTAACAGATCTATTAAAGATTA 660
Qy	221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db	661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATACAAAGAAAAATAACTACTACT 720
Qy	241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db	721 AAGATATACTCATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAAATTT 780
Qy	261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db	781 AAAAAAATAAAAAATAAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy	281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db	841 CCTAGTGATATAAAGCATTTATTTCAAGGAAAA 873
RESULT 5	
AR689942	AR689942 873 bp DNA linear PAT 12-SEP-2005
LOCUS	
DEFINITION	Sequence 11 from patent US 6905867.
ACCESSION	AR689942
VERSION	AR689942.1 GI:74471950
KEYWORDS	

```

QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAATTTTCAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 6
AR691838
LOCUS AR691838 873 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 11 from patent US 6911337.
ACCESSION AR691838
VERSION AR691838.1 GI:74479549
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding beta -1,4-GalNAC transferase
JOURNAL Patent: US 6911337-A 11 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source
Location/Qualifiers
1. .873
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,09e-191 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AR691838 (1-873)
QY 1 MetLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTCGATGTTAGTAGTAATCAATTTTATTGAAAGATAAAATACTAT 120
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAACAGTGTTTTACACCCCTGGTTCTTCTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGCCGAACTAATTAATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTTTAAACAACTTAAGAAATTTAATGCTTTATTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAGAAATTTATTTTCAATCAAGAAATTAACCTTCAGGGGTCTATATGTCAGTAGCC 420

```

```

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTATGATCCAAACCAAGAAATCTTTTAAACCTAGCCCTTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAAATATATTTGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 7
AF215659 11474 bp DNA linear BCT 02-JAN-2001
LOCUS Campylobacter jejuni lipooligosaccharide biosynthesis locus,
DEFINITION partial sequence.
ACCESSION AF215659
VERSION AF215659.1 GI:12004275
KEYWORDS
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.
TITLE The LOS biosynthesis locus of Campylobacter jejuni O:4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 11474)
AUTHORS Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1999) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
source
Location/Qualifiers
1. .11474
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43432"
/serotype="O:4"
/db_xref="ATCC:43432"
/db_xref="taxon:197"
<1. .360
/notes="ORF-01; similar to RfaC from Helicobacter pylori"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAG43973.1"
/db_xref="GI:12004276"
/translations="KEYFAKEVLNLSGIDETHIELAPKFSLEELMAFTKMDLIIGN
DSGPTHLAFLANKASITIFGATPSYRNAFTINKITIDAGKKIQNAKHIDKSDFCITR
IIEEDIFKLAKGLNEK"
350. .1237
CDS
CDS

```

```

/note="ORF-02; similar to Waam from Helicobacter pylori"
/codon_start=1
/transl_table=11
/product="acyltransferase"
/protein_id="AAG43974.1"
/db_xref="GI:12004277"
/translation="MNSDRIYLSLYILKFPVTPMPCILHFLALIVARITFVILNKK
HRKIIDNLQICFPYOTQKRDKLSKIYENPAQGIQCLQNTTKKILNKVNFN
ENLLDALAKRPPIITTAHYGNWELLSAYAKGALSIVGKKUKSEIYELISQNR
TQFDIELDKGGIRQMSALKRGALGILTDQCVKESVRLKFPFNKEVNYQMGASL
IAQRNALIIPVAYKEGGKFCIEFFKAKDSQNASLEBLTYQAQSCBEMIKRPWEY
FFHRRPASVNEEIIYKGA"
1234. 2490
/note="ORF-03; similar to LgtF from Neisseria
meningitidis"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAG43975.1"
/db_xref="GI:12004278"
/translation="MNLKQISVIIIVKNAQTLLECLNSLKDQFDIIILLNSESINTL
NIANKFKDFANFYIYHNAFTGFGALKNLASYAKDNWILSDADAEVLENCIKELN
LKQDNIIALSRKLYKGEMIKACGMPDYVLRIFNKNFTFNDNLVHESVLPLSKNA
KKIYLKGLKHYSDKISHLIDKMOYSSSLWAKONIHKKSGVLKANLAFWFFRNYF
LKNGFLYGGKGIISVCALGTFPKYMKLYELQROKPTCALIITYNQKELKLALVD
SVKJLAFUPNEIADDSKEDTARLIEYQKDFCPDLKHQWQDEGFKLSRKNKI
KNADSEYIIVDGMIDLEKOFIHEHFAQRKFLQGSRVILNKKESSELNKKDYRI
IPNKKDFKSKNSFLAKIFYLSKRR"
2786. 3955
/note="ORF-04; local similarity (N-terminus) to Cps14J
from Streptococcus pneumoniae"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAG43976.1"
/db_xref="GI:12004279"
/translation="MKKIIGVPIYNVKYLRECLDSVINQTYTNLEIILVNDGSTE
HSLNIAKYLTKDKRITLFDKNGKNSGARNIGFYSGEYKLNKNTQHKNSLIEP
QLDGNPNYIYKAYSQSAFNEKDLTFTYPSIDYIIFLDSNWKLNCEBECVIRM
KNVDVFDHDCITDNDIKNHKKTRMEIFDFKCEIITPKEYNANLALSGVSRDISFG
WNGMIDNFLQYIKLKFNIINEDIHFGIILFASAKIYVLSQKLYLCRLRANSISN
HDKITKANVSFKDIIVETFCENAKAYLKAASRVITALKLIEFFKDKQKNALAL
IKETFLPCYAKALMIKKPKDPLNLKEQLVLIKPFIQTKLPYDIWKFWQIKNI"
4025. 5068
/gene="CgtA"
4025. 5068
/gene="CgtA"
/note="glycosyltransferase; similar to CgtA from
Campylobacter jejuni OH4384 encoded by GenBank Accession
Number AF130984"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylglactosaminyltransferase"
/db_xref="GI:12004280"
/translation="MLFOSYFVKIICLFIPPRKIRHKIKKTFLLKNIQBDKIDSVLPK
KTLVQINYNEDLILKNALIREHGKGYFNDEKSDPKSLNPFATVRKNEAITL
KASLESILPAIORGVIGNDCITDGEIILEFCQYPSFIPIKYPIEIOIQNPKEEN
KLYSYNVVASFIPKDEWLIKIDVDHIYDAKKLYKSYFIPKNKYDVVSYSRYDIHYFN
DNFIFCKDNNGNILEPQDCLINYNLKWKEVLIDRINNWKKATQKQSFSSNIHSL
QUKYKRIILFHTLENNYHFFPUKKRAQDIYKYNWISIEEFKFFLQINHNHIEFSMI
SKETLKIKFLTLF"
complement(5054..5959)
/gene="CgtB"
complement(5054..5959)
/gene="CgtB"
/note="glycosyltransferase; similar to CgtB from
Campylobacter jejuni OH4384 encoded by GenBank Accession
Number AF130984"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAG43978.1"

```

```

/db_xref="GI:12004281"
/translation="MFKISIIILPTYNVEQYIARAIESCINOTFKDIEIIVVDCGNDN
SINIAREYKDKRKIKLIHNEKNLGLLRARYSGVKVANSPIYMLDPDDDLNACBE
CIKILDEQDEVLVFFNAIVESNVISYKFKDFNSGFGYSKEFVKKILAKKNLYWTMG
KLIIRKKLYLEAFASLKEKDKVKNMAEDVLLYYPMLSQOAKIAYMNCNLVHYVPNNNS
ICNTNEVLVKNNIQELQVLNLYLRQYILNKYSVLYVLIRKLYIYQIYIKIRTKLM
VTLAKINILITKLILFKYKKFLKQC"
6048. 6923
/gene="cst-II"
6048. 6923
/gene="cst-II"
/note="sialyltransferase; activity determined
experimentally"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAG43979.1"
/db_xref="GI:12004282"
/translation="MKNVILIAQNGPSLKEIDYSRLPNDVDFRCNOFYFEDKYVLGKG
CAKAVTPGCFPEQYVYTLKLIQNOBYETELIMCSNNOAHLENENFVKTFYDFPDA
HLGYDFKQLKEFNAYFKEHEIYFNQRIISGVYMCATAIAGKELYLSGIDFVONGS
STAFDTKQENLLKAPDFKNDKSHYIGHKNDIKALEFLEKTYKIKLYCLCPNSLLA
NFIELAPNLNSNFIIOEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENVYVYKLIKD
LURLPSDIKHYPKKG"
6924. 7964
/note="ORF-08; similar to SiaC from Neisseria
meningitidis"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAG43980.1"
/db_xref="GI:12004283"
/translation="MKEIKQNIISSEKAPLVVPEIGINHNGSLELAKIMVDAFSA
GAKIKHQTHIVEDENSKAAKVIPOGNAKISYIIEIMQKALNTKDELALKEYTEKGL
VYLSTPFSRAGVNELEDGVSFAFKIGSGECNNYPLIKHAAFKKPMIVSGMNSIESI
KPTVKILLDNEIPEFVLMHTNLYPTPHNLVRLNMLEKFEFCMVLGSDHTDNLAC
LGAVALGACLERHFTDSMHRSGPDIIVCSMDTOALKELIIQSQMAMRGNNSKKA
KQEQVITDAPASVWSIKDKKGEVLSMDNIWVKRPLGLGISASEFENILGKALRDI
ENDTQUSYEDFI"
8021. 9079
/note="ORF-09; similar to SiaA from Neisseria
meningitidis"
/codon_start=1
/transl_table=11
/product="sialic acid biosynthesis protein"
/protein_id="AAG43981.1"
/db_xref="GI:12004284"
/translation="MYRVQNSSEFELYIFATGMHLSKNFGYTVKELYKNGFKNIYEFI
NYDKYQTKALATTIDGFSRYVNELKPDILVHGDRIEPLAAAIAGALNVLVNDH
GSEISTIDDSLRHAIKSLAHILVNDEFAKRELMOLGEDEKSIIFIGSPDELELND
KISLSEAKYIDINENYALLMFHPVTEITSKNQADNLVKKALIQSNKRYIVIPNN
DIGFELILOSIEELKNRPFPSPURFYFITLLKNADFIIGNSSCILKEALYKTA
GILVRSRQNGRLGENTLKNVANSDEILKAINTIHKKQDLFSAKLEILDSKLPFEYL
LSGEFFKLTQKVFQDIK"
9076. 9741
/note="ORF-10; similar to NeuA from Haemophilus ducreyi"
/codon_start=1
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AAG43982.1"
/db_xref="GI:12004285"
/translation="MSLAIIPARGSGKIKKNKVLNKNKPLIYVYTIKAALNTKSISK
VYVSDSEILNAYKSNVDILKRPISLAQDDTTSKQVLLHALFKYKIDVDFVLPQT
SPURTNIHIDEAFNLYKSNALISVQCDNKILKAFVCDNGDLAGIENDEYFPI
ROKLPKTYMSNGAIYILKIKEFLNPSFLQNTKTHFLMDESSLDIDCLELKKVEQI
WKK"
9729. 10562
/note="ORF-11; similar to putative ORF from Rhodobacter
capsulatus encoded by GenBank Accession Number AF010496"
/codon_start=1
/transl_table=11
/product="acetyltransferase"
/protein_id="AAG43983.1"

```



```
/db_xref="GI:12004286"
PLENNQLESVAGLGEIEEYVAFKFSNLHEMGSPFSQFLPHYAKVGRYCSADGVSM
ENFQPMRIRISTASTYETNHSFINNACONHINKTPPIVNHNPSSSIHLIILODDWMI
GKDVLLKGQITLGTGCVIGQRAVTVNDVPPYAIIVAGIPAKIIRKIFRNKKTIERLKIQ
WMKYHFADFYIDLKINQYUJDLLEEKIIRKSI SYNPKNKLYFRDLLELKSKKIPNL
F"
CDS
Complement (10554..11366)
/note="ORF-12; similar to putative ORF HI0868 from
Haemophilus influenzae encoded by GenBank Accession Number
U32768"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAG43984.1"

Alignment Scores:
Pred. No.:      3.94e-190      Length:      11474
Score:          1563.00      Matches:      291
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      100.0%      Indels:      0
DB:              15          Gaps:          0

US-10-734-719-12 (1-291) x AF215659 (1-11474)
QY      1  MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      6048  ATCAAAAAAGTTATTTCCTGCGAAATGACCAAGTTTAAAAAGAAATTGATTATTCACAGG 6107
QY      21  LeuProAsnAspPheAspValPheArgCysAenGlnPheTyrPheGluAspLysTyrTyr 40
Db      6108  CTACCAAAATGATTTCATGATGATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
QY      41  LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      6168  CTTGTAAAAAATGCAAAGCAGTGTTTACACCCCTGGTCTCTCTTTGAGCAATACTAC 6227
QY      61  ThrLeuLysHisLeuIleGlnAenGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      6228  ACTTTAAAAACATTAATCCAAAATCAAGAATATGAGACCGCAACTAATTATGTTCTTAAT 6287
QY      81  TyrAenGlnAlaHisLeuGluAenGluAenPheValLysThrPheTyrAspTyrPhePro 100
Db      6288  TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACCAGTATTTTCTCT 6347
QY      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      6348  GATGCTCATTTGGGATATGATTTTTTTAAACCACTTAAGAATTTAATGCTTATTTTAAA 6407
QY      121  PheHisGluIleTyrPheAenGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      6408  TTTTCAGAAATTTATTTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 6467
QY      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAenGlySer 160
Db      6468  ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAATTTGATTTTATCAAAATGGGTCA 6527
QY      161  SerTyrAlaPheAspThrLysGlnGluAenLeuLeuLysLeuAlaProAspPheLysAsn 180
Db      6528  TCTTATGCTTTTGTATACCAACAAGAAAATCTTTTAAAACAGCCCTGATTTTAAAAAT 6587
QY      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      6588  GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTAGATTTTCTA 6647
QY      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      6648  GAAAAAACTTACAAAAATAAACTATATTCGTTATGTCCTAACAGTCTTTTAGCAATTTT 6707
QY      221  IleGluLeuAlaProAsnLeuAenSerAsnPheIleIleGlnGlyLysAsnTyrThr 240
Db      6708  ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 6767
```

```
QY      241  LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      6768  AAAGATATACTCATACCTTCTAGTGAGCTTATCGAAAAATTTTCAAAAAATTAATAATTT 6827
QY      261  LysLysIleLysIleLysGluAenValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      6828  AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAGATTA 6887
QY      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      6888  CCTAGTATATAAAGCATTATTTCAAAAGGAAAA 6920

RESULT 8
AX934431
LOCUS      AX934431      876 bp      DNA      linear      PAT 05-JAN-2004
DEFINITION Sequence 8 from Patent WO02074942.
ACCESSION  AX934431
VERSION     AX934431.1  GI:40641678
KEYWORDS
SOURCE      Campylobacter jejuni
ORGANISM    Campylobacter jejuni
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
            Campylobacteraceae; Campylobacter.
REFERENCE   1
AUTHORS     Gilbert M. and Wakarchuk W.W.
TITLE       Campylobacter glycosyltransferases for biosynthesis of gangliosides
            and ganglioside mimics
JOURNAL     Patent: WO 02074942-A 8 26-SEP-2002;
            National Research Council of Canada (CA)
FEATURES    Location/Qualifiers
             1..876
             /organism="Campylobacter jejuni"
             /mol_type="unassigned DNA"
             /db_xref="taxon:197"
             1..876
             /notes="unnamed protein product; Campylobacter
             alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.
             jejuni O:19"
             /transl_table=11
             /codon_start=1
             /protein_id="CAFO4169.1"
             /db_xref="GI:40641679"
             /translations="MKKVIITAGNGPSLKEIDYSRLPNDFDVFRGNQFYEDKYLGKK
             KCAVFTPNPFQYQYTLKHLIQNOEYETELINGSNYNQAHLENFENFKTVDFPDPA
             HLGYDFKQLKEFNAYPKHEIIFYNQRTSGVYMCVAIALGYKEIYLSGLDFYQNGS
             SYAFDFKQENLLKLPDFKNDRSHYIGHSKNTDIKALEFLEKTKIKLKYLCPCNSLLA
             NFTELAPNLSNFIIQEKNNYTKDILIPSSSEAYGKFSKNINFKKIKIKENVYVKLIKD
             LLRLPSDIKHVFKGK"

ORIGIN
Alignment Scores:
Pred. No.:      1.84e-190      Length:      876
Score:          1557.00      Matches:      290
Percent Similarity: 99.7%      Conservative: 0
Best Local Similarity: 99.7%      Mismatches: 1
Query Match:      99.6%      Indels:      0
DB:              2          Gaps:          0

US-10-734-719-12 (1-291) x AX934431 (1-876)
QY      1  MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1  ATGAAAAAAGCTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCACAGG 60
QY      21  LeuProAsnAspPheAspValPheArgCysAenGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAATGATTTCATGATGATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY      41  LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      121  CTTGTAAAAAATGCAAAAGCAGTCTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180
```

```
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGCGCACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAenPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAGAAATCTTTTACGATTATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgileThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCCTCGCATATCGACATAGTAAATATAGATATCAAGATATAAAGCTTTAGAAATTTCT 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCAACAAATTAACATATATGCTTATGCTTAATAGTCTTTTACCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTTTAAATTTTATCATAAGAAATTAATTAATTTT 720
QY 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGAAATTTTCAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTAACAGTTGATAAAGATCTATTAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 9
BD249794
LOCUS
DEFINITION
Campylobacter glycosyltransferases for biosynthesis of
gangliosides and ganglioside mimics.
ACCESSION
BD249794.1 GI:33059564
VERSION
JP 2002535992-A/5.
KEYWORDS
Campylobacter jejuni
ORGANISM
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 876)
AUTHORS
Gilbert,M. and Wakarchuk,W.W.
TITLE
Campylobacter glycosyltransferases for biosynthesis of
gangliosides and ganglioside mimics
JOURNAL
Patent: JP 2002535992-A 5 29-OCT-2002;
NATIONAL RESEARCH COUNCIL OF CANADA
COMMENT
OS Campylobacter jejuni
PN JP 2002535992-A/5
PD 29-OCT-2002
PF 01-FEB-2000 JP 2000597438
PR 01-FEB-1999 US 60/118213.31 -JAN-2000 US 06/495406 PI
MICHEL GILBERT, WARREN W WAKARCHUK
PC C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC
```

```
Campylobacter glycosyltransferases for biosynthesis of CC
gangliosides and
PH Key Location/Qualifiers
FT source 1..876
FT /organism='Campylobacter jejuni'.
FEATURES
source
1..876
/organism='Campylobacter jejuni'
/mol_type='genomic DNA'
/db_xref='taxon:197'
ORIGIN
Alignment Scores:
Pred. No.: 1,84e-190 Length: 876
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 2 Gaps: 0
US-10-734-719-12 (1-291) x BD249794 (1-876)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATCAAAAAAGTTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTTCATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAATAATGCAAGCAGTGTTTACCCCTTAATTTCTTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGCACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgileThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCATATATCGACATAGTAAATATAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCAACAAATAAACTATATTTGCTTATGCTTAATAGTCTTTTACGAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTTTAAATTTTATCATAAGAAATTAATTAATTTT 720
QY 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
```

```

Db      721  AAAGATATACTATACATCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy      261  LysLysLysLysLysLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeuArgLeu 280
Db      781  AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841  CCTAGTGATATAAAGCATTTATTTCAAGGAAAA 873

RESULT 10
LOCUS   AR271703              876 bp      DNA      linear      PAT 10-APR-2003
DEFINITION Sequence 8 from patent US 6503744.
ACCESSION AR271703
VERSION   AR271703.1 GI:29703248
KEYWORDS .
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS  Gilbert,M. and Wakarchuk,W.W.
TITLE    Campylobacter glycosyltransferases for biosynthesis of gangliosides
          and ganglioside mimics
JOURNAL  Patent: US 6503744-A 8 07-JAN-2003;
          National Research Council of Canada, Ottawa;
          CAX;

FEATURES             Location/Qualifiers
     source           1..876
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:          1..84e-190      Length:          876
Score:              1557.00          Matches:          290
Percent Similarity: 99.7%            Conservative:      0
Best Local Similarity: 99.7%          Mismatches:       1
Query Match:        99.6%            Indels:           0
DB:                 2                Gaps:             0

US-10-734-719-12 (1-291) x AR271703 (1-876)

Qy      1  MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1  ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAATGATTGATGATTATTAGATGAATCAATTTTATTGGAAGATAAATACTAT 120
Qy      41  LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      121  CTGGTAAAAAATGCAAGACAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
Qy      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181  ACITTTAAACATTTAATCCAAAATCAAGAATATGAGACCGCAACTAATATTATGTGTTCTAAT 240
Qy      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATATTATTTCCCT 300
Qy      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAA 360
Qy      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTTCAAGAAATTTATTTCATCAAGAAATTAACCTCAGGGCTTATATGTGTCAGTAGCC 420
Qy      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

```

```

Db      421  ATAGCCCTAGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db      481  TCTTATGCTTTTGNATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTTACAAAAATAAACTATATATGCTTATGCTCTCTAATAGTCTTTTACCAATTTT 660
Qy      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db      661  ATAGAAGTACGCCCAATTTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 720
Qy      241  LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721  AAAGATATACTCATACCTTCTAGTGAGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
Qy      261  LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781  AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841  CCTAGTGATATAAAGCATTTATTTCAAGGAAAA 873

RESULT 11
LOCUS   AR481785              876 bp      DNA      linear      PAT 14-MAY-2004
DEFINITION Sequence 8 from patent US 6699705.
ACCESSION AR481785
VERSION   AR481785.1 GI:47243420
KEYWORDS .
SOURCE   Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS  Gilbert,M. and Wakarchuk,W.W.
TITLE    Campylobacter glycosyltransferases for biosynthesis of gangliosides
          and ganglioside mimics
JOURNAL  Patent: US 6699705-A 8 02-MAR-2004;
          National Research Council of Canada, Ottawa;
          CAX;

FEATURES             Location/Qualifiers
     source           1..876
                     /organism="unknown"
                     /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:          1..84e-190      Length:          876
Score:              1557.00          Matches:          290
Percent Similarity: 99.7%            Conservative:      0
Best Local Similarity: 99.7%          Mismatches:       1
Query Match:        99.6%            Indels:           0
DB:                 2                Gaps:             0

US-10-734-719-12 (1-291) x AR481785 (1-876)

Qy      1  MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1  ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAATGATTGATGATTATTAGATGAATCAATTTTATTGGAAGATAAATACTAT 120
Qy      41  LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      121  CTGGTAAAAAATGCAAGACAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
Qy      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181  ACITTTAAACATTTAATCCAAAATCAAGAATATGAGACCGCAACTAATATTATGTGTTCTAAT 240
Qy      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATATTATTTCCCT 300
Qy      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAA 360
Qy      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTTCAAGAAATTTATTTCATCAAGAAATTAACCTCAGGGCTTATATGTGTCAGTAGCC 420
Qy      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

```

```
Db 121 CTTGGTAAAAAATGCAAGACAGTGTCTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCACTAGAAAATGAAAATTTGTGAAAACTTTTACGATTTATTTCTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAAATTTATTTCAATCAAGAATTAATCTCAGGGGTCTATATGTCGAGTAGGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAGAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGTATACCAACAAAGAAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGATATGATTTTAAACAACTATACAGATATAAAGCTTTTGA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTAAAGAAAATGTTTATTTACAGTTGATTAAGAACTATTAAGATTATTA 780
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTCATACCTCTTAGTAGGGCTTATGGAATAATTTTCAAAAAATTTAAT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 781 AAAAAAATTTAAAGAAAATGTTTATTTACAGTTGATTAAGAACTATTAAGATTATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
```

RESULT 12  
AR527384  
LOCUS AR527384 876 bp DNA linear PAT 08-OCT-2004

DEFINITION Sequence 8 from patent US 6723545.

ACCESSION AR527384

VERSION AR527384.1 GI:53914405

KEYWORDS

SOURCE

ORGANISM

Unclassified.

1 (bases 1 to 876)

AUTHORS Gilbert,M. and Wakarchuk,W.W.

TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity

JOURNAL Patent: US 6723545-A 8 20-APR-2004;

National Research Council of Canada; Ottawa;

CAX;

FEATURES

source

1..876

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 1.84e-190 Length: 876  
Score: 1557.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.6% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AR527384 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20

Db 1 ATGAAAAAGTATTATTCTGGAATGGACCAAGTTAAAGAAATTTGATTTATTTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

Db 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60

Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80

Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATTTATGTGTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100

Db 241 TACAACCAAGCTCACTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

Db 361 TTTTCAGAAAATTTATTTCAATCAAGAATTAATCTCAGGGGTCTATATGTCGAGTAGGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATAGCCCTAGGATACAGAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180

Db 481 TCTTATGCTTTTGTATACCAACAAAGAAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

Db 541 GATCGCTCGCACTATATCGGATATGATTTTAAACAACTATACAGATATAAAGCTTTTAGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220

Db 601 GAAAAAATTTAAAGAAAATGTTTATTTACAGTTGATTAAGAACTATTAAGATTATTA 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240

Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACTACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

Db 721 AAAGATATATCTCATACCTCTTAGTAGGGCTTATGGAATAATTTTCAAAAAATTTAAT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280

Db 781 AAAAAAATTTAAAGAAAATGTTTATTTACAGTTGATTAAGAACTATTAAGATTATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291

Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 13

AR609663

LOCUS

DEFINITION

Sequence 8 from patent US 6825019.

AR609663 876 bp DNA linear

Sequence 8 from patent US 6825019.

```
ACCESSION AR609663 GI:56664963
VERSION AR609663.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert M. and Wakarchuk, W.W.
TITLE Polypeptides having beta-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 8 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
    source
        1..876
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
  Pred. No.: 1.84e-190 Length: 876
  Score: 1557.00 Matches: 290
  Percent Similarity: 99.7% Conservative: 0
  Best Local Similarity: 99.7% Mismatches: 1
  Query Match: 99.6% Indels: 0
  DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AR609663 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCCTGGAAATGACCAAGTTAAAGAAATTCATTATTCACGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAAAATGATTTGATGTAATCAATTTTATTTTGAAGATAATATCTAT 120

QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyr 60
DB 121 CTGGTAAAAATGCAAGACAGTGTTCACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTATTCACAAATCAAGATATGAGACCGAATATTTATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAATTTGTAAAACTTTTACGATTTATTTTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCAGCAAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTCGACGATGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCGACATATTCGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTACGAAATTT 660

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240

Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATCAAGAAAAAATACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTCATACCTCTTAGTGAGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 14
AR689941
LOCUS AR689941
DEFINITION Sequence 8 from patent US 6905867.
ACCESSION AR689941
VERSION AR689941.1
KEYWORDS GI:74471949
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Nucleic acids encoding polypeptides with beta.1,3-galactosyl transferase activity
JOURNAL Patent: US 6905867-A 8 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
    source
        1..876
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
  Pred. No.: 1.84e-190 Length: 876
  Score: 1557.00 Matches: 290
  Percent Similarity: 99.7% Conservative: 0
  Best Local Similarity: 99.7% Mismatches: 1
  Query Match: 99.6% Indels: 0
  DB: 2 Gaps: 0

US-10-734-719-12 (1-291) x AR689941 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCCTGGAAATGACCAAGTTAAAGAAATTCATTATTCACGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAAAATGATTTGATGTAATCAATTTTATTTTGAAGATAATATCTAT 120

QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyr 60
DB 121 CTGGTAAAAATGCAAGACAGTGTTCACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTATTCACAAATCAAGATATGAGACCGAATATTTATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAATTTGTAAAACTTTTACGATTTATTTTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCAGCAAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTCGACGATGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCGACATATTCGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTACGAAATTT 660

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240
```



GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 14:18:17 ; Search time 519 Seconds  
(without alignments)  
5863.943 Million cell updates/sec

Title: US-10-734-719-12

Perfect score: 1563

Sequence: 1 MKKVIIAGNPSLKEIDYSR.....KLIKDLLRLPSDKHFKGK 291

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/abss/ABSSWEB spool/US10734719/runat\_21082006\_141806\_12305/app\_query.fasta\_1  
-DB=N\_Geneseq -QFMT=fastap -SUFFIX=P2N.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05p  
-USER=US10734719 @CGN\_1\_1\_1423 @runat\_21082006\_141806\_12305 -NCPU=6 -ICPU=3  
-NO\_WMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_8.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*  
15: Geneseqn2006s.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1563	100.0	873	8	ABT13670 C. jejuni
2	1557	99.6	876	3	AA53726 Campyloba
3	1557	99.6	876	8	ABT13669 C. jejuni

4	1553	99.4	873	8	ABT13671
5	1534	98.1	876	3	AA53724
6	1534	98.1	876	8	ABT13667
7	1531	98.0	876	13	ADU77566
8	1531	98.0	12219	13	ADU77595
9	1524	97.5	876	3	AA53721
10	1524	97.5	876	8	ABT13666
11	1524	97.5	11474	3	AA53720
12	1524	97.5	11474	8	ABT13665
13	1513	96.8	876	3	AA53725
14	1513	96.8	876	8	ABT13668
15	749	47.9	1293	2	AAZ25693
16	743.5	47.6	19196	13	ADT05647
17	729.5	46.7	13379	13	ADT05530
18	729.5	46.7	117576	13	ADT05736
19	729.5	46.7	349980	13	ADT05648
20	720.5	46.1	110000	2	AACT42063_03
21	698.5	44.7	909	8	ACC71694
22	691.5	44.2	4277	13	ADT05421
23	120	7.7	6128	6	ABQ67039
24	120	7.7	6129	6	ABL70537
25	120	7.7	6129	6	ABL70537
26	117.5	7.5	1866	2	AA53720
27	116	7.4	4838	2	AAV70891
28	115	7.4	110000	6	ABA92787_2
29	114	7.3	110000	2	AAZ20248_06
30	114	7.3	111309	2	AAZ20250
31	112.5	7.2	110000	2	AAZ20250
32	112	7.2	4810	2	AAV70887
33	112	7.2	11474	8	ABT13665
34	112	7.2	11474	8	ABT13665
35	112	7.2	50000	6	ABL55643
36	112	7.2	50000	6	ABL55643
37	110.5	7.1	186854	14	ADX38909
38	110	7.0	1179	3	AA53725
39	110	7.0	110000	2	AAZ20248_06
40	109.5	7.0	2727	8	ACA23131
41	109.5	7.0	3366	3	AA53725
42	109.5	7.0	7758	6	ABL33102
43	109.5	7.0	8365	2	AAZ20056
44	109.5	7.0	110000	2	AAZ20248_03
45	109	7.0	4197	3	AAA70227

ALIGNMENTS

RESULT 1

ABT13670

ID ABT13670 standard; DNA; 873 BP.

XX AC ABT13670;

XX DT 07-FEB-2003 (first entry)

XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #5.

XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
XX KW GalNAc transferase; N-Acetylglucosamine transferase;  
XX KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
XX KW cytidine 5'-monophosphate sialic acid synthetase;  
XX KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
XX KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.

XX PN WC200274942-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-CA000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX XX

PA (CANA ) NAT RES COUNCIL CANADA.  
XX Gilbert M, Wakarchuk WW;  
XX WPI; 2003-040554/03.  
DR P-PSDB; ABU18484.  
XX  
PT New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
PS Disclosure; Page 98-99; 107pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
CC Acetylglucosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention  
XX  
SQ Sequence 873 BP; 350 A; 118 C; 110 G; 295 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5,61e-153 Length: 873  
Score: 1563.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
Gaps: 0

US-10-734-719-12 (1-291) x ABF13670 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGTTAAAGAAATTCATTATCAAGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAAAATGATTTTCATGATTAATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
DB 121 CTTGGTAAATAATGCAAGACATGTTTACACCCCTGGTTCCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCCACTAATATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAAAAATTTTACGATTATTTTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCAVTTGGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTATTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTTCAGAAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTCGATGACC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180

DB 481 TCTTAATGCTTTTGTATACCAACAAGAAAAATCTTTAAAAACTAGCCCTGATTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTCAAAATAAAATCTATATGCTTATGCTTACAGCTCTTTTAGCAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
DB 661 ATAGAACCTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAACTACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATATCTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAGATTTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 2  
AAA53726  
ID AAA53726 standard; DNA; 876 BP.  
XX  
AC AAA53726;  
XX  
DT 15-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE  
XX  
XX  
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside; ds.  
XX  
OS Campylobacter jejuni; O:19 serotype.  
XX  
FH Location/Qualifiers  
CDS 1..876  
FT /\*tag= a  
FT /product= "CstII sialyltransferase"  
XX  
PN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CA000086.  
XX  
PR 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX WPI; 2000-524418/47.  
DR P-PSDB; AAY97212.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
PS Disclosure; Page 96; 120pp; English.  
XX



CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesizing sialylated oligosaccharides such as ganglioside,  
 CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyze the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 15-SEP-2003 to standardise OS field)  
 CC  
 XX  
 SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,37e-152 Length: 876  
 Score: 1557.00 Matches: 290  
 Percent Similarity: 99.7% Conservative: 0  
 Best Local Similarity: 99.7% Mismatches: 1  
 Query Match: 99.6% Indels: 0  
 DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x AAA53726 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 DB 1 ATGAAAAAGTTATTATTCTGGAATGACCACAAAGTTTAAAGAAATTTGATTATTCAAGG 60  
 QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
 DB 61 CTACCAATGATTATTGATGTTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
 QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
 DB 121 CTTGGTAAAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180  
 QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
 DB 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGACCGAATTAATTATGTGTTCTAAT 240  
 QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAACAACTTTTACGATTATTTTCCT 300  
 QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 DB 301 GATGCTCATTTGGGATATGATTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
 QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 DB 361 TTTTCAGAAATTTATTTCATCAAGAATTTACCTTCAGGGGTCTATATGTGTCAGTAGCC 420  
 QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 DB 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480  
 QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
 DB 481 TCTTATGCTTTTGTATACCAACAAGAAATCTTTTAAACACTAGCCCTCGATTTTAAAT 540  
 QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
 QY 201 GluLysThrTyrLysLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 DB 601 GAAAAAATTTACAAAATPAAACATATATGCTTATGCTCTAATAGTCTTTTACAAATTTT 560  
 QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
 DB 661 ATAGAACTAGCGCAAAATTTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720  
 QY 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 DB 721 AAAGATATACTCATACCTTCTAGTAGGCTTATGAAAAATTTTCAAAAAATATTAAATTTT 780  
 QY 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
 DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 840  
 QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 DB 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

## RESULT 3

ABT13669  
 ID ABT13669 standard; DNA; 876 BP.

XX AC ABT13669;  
 XX DT 07-FEB-2003 (first entry)  
 XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #4.  
 XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX OS Campylobacter jejuni.

XX WO200274942-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-CA000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX PA (CANA ) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX WP1: 2003-040554/03.

XX P-PSDB; ABJ18482.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.

XX PS Disclosure; Page 98; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune

CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention

XX SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.37e-152 Length: 876  
Score: 1557.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.6% Indels: 0  
DB: 8 Gaps: 0

US-10-734-719-12 (1-291) x ABT13669 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGACCACAGTTTAAAGAAATTTGATTATTCAAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTGAGATAAATACTAT 120  
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
DB 121 CTGTGTAATAAATGCAAGACGATGTTTTACACCCCTAATTTCTTTGAGCAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACAATTTAATCAAAATCAGAAATATGAGCCGAACATAATTATGTGTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGAAAACTTTTACGATTATTTTCCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACATTAAGAATTTAATGCTTTATTTTAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTCAGAAATTTATTTCAATCAAGAATTTACTCTCAGGGGCTATATGTGTCAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTTATGCTTTTGTATACCAACAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAA 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCGCACTATTCGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAATCAAAAATAAAATATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACAC 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAAATAAATAAATAAGAAAATGTTTATTATACAGTTGATTAAGATCTATTAGATTA 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 4

ABT13671  
ID ABT13671 standard; DNA; 873 BP.

AC ABT13671;

DT 07-FEB-2003 (first entry)

DE C. jejuni bifunctional sialtransferase cstII coding sequence #6.

XX Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylglucosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.

XX Campylobacter jejuni.

XX WO200274942-A2.

XX 26-SEP-2002.

PF 22-FEB-2002; 2002WO-CA000229.

XX 21-MAR-2001; 2001US-00816028.

XX (CANA ) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2003-040554/03.

DR P-PSDB; ABU18485.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.

XX Disclosure; Page 99; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
CC Acetylglucosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention

XX Sequence 873 BP; 353 A; 117 C; 107 G; 296 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 6.16e-152 Length: 873  
Score: 1553.00 Matches: 289  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 2  
Query Match: 99.4% Indels: 0  
DB: 8 Gaps: 0

US-10-734-719-12 (1-291) x ABT13671 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
|||||

```
Db      1  ATCAAAAAAGTTATTATTCTGCTGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
Qy      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAATGATTTTGTATGTTAGATGTAATCAATTTTATTTTTGAAGATAAATACTAT 120
Qy      41  LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      121  CTTGGTAAAAATGCAAAACAGTGTTTACACCCCTAATTTCTCTTCCTGAGCAATACTAC 180
Qy      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181  ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATTATGTGTTCTAAT 240
Qy      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAAAAAGCTTTTACGATTATTTCCT 300
Qy      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATGATGATTTTAAAAACAACTTAAGAATTTAATGCTTATTTTAAA 360
Qy      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTCAGGAAATTTATTTCAATCAAGAATTACTCAGGGGTCTATATGTGTGAGTAGCC 420
Qy      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db      481  TCTTATGCTTTTTCATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTTACAAAATAAACTATATTGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660
Qy      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240
Db      661  ATAGAACTAGCCCAATTTAAATTCAAATTTTATCATACAGAAAAAATAACTACACT 720
Qy      241  LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721  AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAAATTTTCAAAAAATATTATTTT 780
Qy      261  LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781  AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGACTCTATTAAAGATTA 840
Qy      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841  CCTAGTGATATAAAGCATTATTTTCAAGGAAAA 873
```

## RESULT 5

AAAS3724

ID AAAS3724 standard; DNA; 876 BP.

XX

AC

AAAS3724;

XX

DT

15-SEP-2003 (revised)

XX

DE Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.

XX

KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;

KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;

KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;

KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;

immunity; immunogen; ganglioside; ds.

XX

OS Campylobacter jejuni; O:10 serotype.

XX

FH Key Location/Qualifiers

FT CDS

1. .876

/\*tag= a

/product= "alpha-2,3-sialyltransferase"

XX

PN W0200046379-A1.

XX

10-AUG-2000.

XX

01-FEB-2000; 2000WO-CA000086.

XX

01-FEB-1999; 99US-0118213P.

XX

31-JAN-2000; 2000US-00495406.

XX

(CAN ) NAT RES COUNCIL CANADA.

XX

Gilbert M, Wakarchuk WW;

XX

WPI; 2000-524418/47.

XX

P-PSDB; AAY97210.

XX

Novel glycosyltransferase polypeptides and polynucleotides useful for

XX

biosynthesis of ganglioside and ganglioside mimics, as diagnostic

XX

reagents and as immunogen for producing antibodies.

XX

Claim 6; Page 92; 120pp; English.

XX

A reaction mixture for the synthesis of a sialylated oligosaccharide is useful for synthesising sialylated oligosaccharide such as ganglioside, lyso ganglioside or their mimics. Glycosyltransferases are useful for chemo-enzymatic synthesis of oligosaccharides, including gangliosides and other oligosaccharides that have biological activity. The enzymes and nucleic acids that encode them are useful for studies of the pathogenesis mechanisms of organisms that synthesize ganglioside mimics, such as C. jejuni and the nucleic acids are used as probes to study expression of genes involved in ganglioside mimetic synthesis. Antibodies raised against the glycosyltransferases are also useful for analyzing the expression patterns of these genes involved in pathogenesis. The nucleic acids are also useful for designing antisense oligonucleotides for inhibiting expression of the Campylobacter enzymes that are involved in the biosynthesis of ganglioside mimics that can mask the pathogens from the host's immune system. The oligosaccharides are useful as diagnosing reagents or as therapeutics and as immunogens for producing antibodies. Bacterial glycosyltransferase can be used to catalyse the formation of oligosaccharides that are identical to the corresponding mammalian structures and are easier and less expensive to produce in large quantity, compared to the mammalian glycosyltransferase. The bacterial origin of the enzymes facilitates expression of large quantities of the enzymes using relatively inexpensive prokaryotic expression systems.

(Updated on 15-SEP-2003 to standardise OS field)

XX

SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	5.88e-150	Length:	876
Score:	1534.00	Matches:	286
Percent Similarity:	98.6%	Conservative:	1
Best Local Similarity:	98.3%	Mismatches:	4
Query Match:	98.1%	Indels:	0
DB:	3	Gaps:	0

US-10-734-719-12 (1-291) x AAA53724 (1-876)

Qy

1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20

Db

1 ATGAAAAAGTTATTATTGCTGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60

Qy

21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

```
Db 61 CTACCAATGATTTTGATGCTATTTAGATCAATCAATTTTATTATTGAAGATAAACTAT 120
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATTCAAAGCAGTATTTTACAACTCTGGTCTTTTGTGAAACAACTACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAGAATATGAGCGCACTAATATGTTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATCTCAATCAAGAATTACTCAGAGCTATATGTCGAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTGGCTCCTGAATTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCCTCACACTATATCGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GlyLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAAGTTACAAAATAAACTATATTGCTTATGCTCTAACAGCTTTTACCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAAGCTAGGCCCAATTTTAAATTTTAAATTTATCATAAGAAAAAATAACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTTCTAGTGAGGCTTATGGAATTTTTCAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATATTATTACAAGTTGATAAAGATCTATTAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATAAAGCATTTATTTCAAAGGAAAA 873
RESULT 6
ABT13667
ID ABT13667 standard; DNA; 876 BP.
XX
AC ABT13667;
XX
DT 07-FEB-2003 (first entry)
DE
XX
XX C. jejuni bifunctional sialtransferase cstII coding sequence #2.
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAC transferase; N-Acetylgalactosamine transferase;
KW Galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
OS Campylobacter jejuni.
```

```
PN WO200274942-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-CA000229.
XX
PR 21-MAR-2001; 2001US-00816028.
XX
PA (CANA ) NAT RES COUNCIL CANADA.
XX
PI Gilbert M, Wakarchuk MW;
XX
DR WPI; 2003-040554/03.
DR P-PSDB; ABJ18480.
XX
PT New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS Claim 8; Page 96-97; 107pp; English.
XX
CC The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present DNA sequence represents a Campylobacter jejuni
CC gene of the invention
XX
SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5-88e-150 Length: 876
Score: 1534.00 Matches: 286
Percent Similarity: 98.6% Conservative: 1
Best Local Similarity: 98.3% Mismatches: 4
Query Match: 98.1% Indels: 0
DB: Gaps: 0

US-10-734-719-12 (1-291) x ABT13667 (1-876)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGATGCTATTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATTCAAAGCAGTATTTTACAACTCTGGTCTTTTGTGAAACAACTACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAGAATATGAGACCGCACTAATATGTTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCACTAGAAATGAAATTTGTAAAACTTTTACGATTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
```

```
Db 361 TTTCCAGAAATTATCTCAATCAAGAAATTTACCTCAGGAGTCTATATGTGTCAGTAGCT 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTGGCTCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGTCTAACAGTCTCTTTTGAAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAGATATATCTATACACTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTTAAGAAAAATATTTATTACAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 7
ID ADU77566 standard; DNA; 876 BP.
XX
AC ADU77566;
DT 10-FEB-2005 (first entry)
XX
DE Sialytransferase (cstII).
XX
KW gene amplification; campylobacter infection; Sialytransferase; cstII; ds.
XX
OS Campylobacter jejuni.
XX
PN WO2004101817-A1.
XX
XX 25-NOV-2004.
XX
XX 19-MAY-2004; 2004WO-AU000654.
XX
XX 19-MAY-2003; 2003AU-00902445.
XX
XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.
XX
XX Shi F, Fry BN, Coloe PJ;
XX
XX WPI; 2004-834013/82.
XX
XX Identifying Campylobacter in a sample to comprises comparing the profile
PT of an amplification product to a comparator Campylobacter species after
PT performing nucleic acid amplification of the wla gene cluster.
XX
XX Claim 6; SEQ ID NO 13; 75pp; English.
PS
XX
XX The invention describes a method of identifying Campylobacter in a sample
CC comprising comparing the profile of an amplification product to a
CC comparator Campylobacter species after performing nucleic acid
CC amplification. Identifying Campylobacter in a sample comprises: (a)
```

```
CC performing nucleic acid amplification by contacting the sample with a
CC pair of nucleic acid primers, where at least one nucleic acid primer is
CC derived from the wla gene cluster of Campylobacter, for a time and under
CC conditions for generation of an amplification product comprising the wla
CC gene cluster of Campylobacter or its portion; (b) preparing a profile of
CC the amplification product; and (c) comparing the profile to a profile of
CC a comparator Campylobacter species, strain or type, and thus, determining
CC the identity of Campylobacter in the sample. Also described are: (a) a
CC pair of nucleic acid primers, where at least one nucleic acid primer is
CC derived from the wla gene cluster of Campylobacter; or (b) a pair of
CC nucleic acid primers for amplification of the wla gene cluster of
CC Campylobacter or its portion; or (c) a probe capable of hybridizing to
CC the wla gene cluster of Campylobacter. The method and kit are useful for
CC identifying and detecting the presence of Campylobacter in a sample, or
CC for diagnosing infections associated with Campylobacter. This sequence
CC represents Sialytransferase, encoded by a gene in the Campylobacter
CC jejuni wla cluster used in the inventive method of detecting
CC Campylobacter jejuni infection.
XX
```

SQ Sequence 876 BP; 356 A; 115 C; 104 G; 301 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.21e-149 Length: 876  
Score: 1531.00 Matches: 285  
Percent Similarity: 98.6% Conservative: 2  
Best Local Similarity: 97.9% Mismatches: 4  
Query Match: 98.0% Indels: 0  
DB: 13 Gaps: 0

US-10-734-719-12 (1-291) x ADU77566 (1-876)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTTCTCGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGTATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATTCAAAGCAGTATTTTACAATCTCTGCTCTTTTTCGAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAATAATATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTTCAGATATTTTCTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAACGAAATTTATCTCAATCAAGAAATTAACCTCAGAGTCTATATGTGTCAGTAGCT 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTGATACCAACCAAGAAATCTTTTAAACTGGCTCTCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuLeuAlaAsnPhe 220
```

Db 601 GAAAAAACTTCAAAAAATAAACTATATTGCTTCTTAACAGCTTTTTCAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCAAATTTTAAATTCAAATTTTATCATACAGAAAAAATAACTACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTAGGAAATTTTCAAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAAAATAAGAAATATTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 8

ADU77595  
ID ADU77595 standard; DNA; 12219 BP.  
AC ADU77595;  
XX  
XX 10-FEB-2005 (first entry)  
XX  
XX Campylobacter jejuni strain ATCC 43469 wla cluster.  
DE  
XX  
XX gene amplification; campylobacter infection; wla; ds.  
KW  
XX  
XX Campylobacter jejuni.  
OS  
XX  
XX WO2004101817-A1.  
PN  
XX  
XX 25-NOV-2004.  
XX  
XX 19-MAY-2004; 2004WO-AU000654.  
XX  
XX 19-MAY-2003; 2003AU-00902445.  
PR  
XX  
XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
PA  
XX  
XX Shi F, Fry BN, Coloe PJ;  
PI  
XX  
XX WPI; 2004-834013/82.  
DR  
XX

Identifying Campylobacter in a sample to comprises comparing the profile of an amplification product to a comparator Campylobacter species after performing nucleic acid amplification of the wla gene cluster.

Disclosure; Fig 6; 75pp; English.

The invention describes a method of identifying Campylobacter in a sample comprising comparing the profile of an amplification product to a comparator Campylobacter species after performing nucleic acid amplification. Identifying Campylobacter in a sample comprises: (a) performing nucleic acid amplification by contacting the sample with a pair of nucleic acid primers, where at least one nucleic acid primer is derived from the wla gene cluster of Campylobacter, for a time and under conditions for generation of an amplification product comprising the wla gene cluster of Campylobacter or its portion; (b) preparing a profile of the amplification product; and (c) comparing the profile to a profile of a comparator Campylobacter species, strain or type, and thus, determining the identity of Campylobacter in the sample. Also described are: (a) a pair of nucleic acid primers, where at least one nucleic acid primer is derived from the wla gene cluster of Campylobacter; or (b) a pair of nucleic acid primers for amplification of the wla gene cluster of Campylobacter or its portion; or (c) a probe capable of hybridizing to the wla gene cluster of Campylobacter. The method and kit are useful for identifying and detecting the presence of Campylobacter in a sample, or for diagnosing infections associated with Campylobacter. This sequence represents a Campylobacter jejuni strain ATCC 43469 wla cluster.

XX

SQ Sequence 12219 BP; 4778 A; 1598 C; 1745 G; 4098 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.44e-148 Length: 12219  
Score: 1531.00 Matches: 285  
Percent Similarity: 98.6% Conservative: 2  
Best Local Similarity: 97.9% Mismatches: 4  
Query Match: 98.0% Indels: 0  
DB: 13 Gaps: 0

US-10-734-719-12 (1-291) x ADU77595 (1-12219)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 5995 ATGAAAAAGTTATTATTTCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 6054  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 6055 CTACCAATGATTTTGATGATTTAGATGCAATTTTATTTTGAAGATAAATACTAT 6114  
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
Db 6115 CTGTGTAATAAATTCAAAGCAGTATTTTACAATCTCTGGTCTTTTGTGAACAACTAC 6174  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 6175 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGACCGAATAATTATGTCTTAAT 6234  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 6235 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCT 6294  
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 6295 GATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAATTTAATGCTTATTTAAA 6354  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 6355 TTTTCAGAAATTTATCTCAATCAAGAAATTTCTCTGGAAATTTTATCAAAATGGGTCA 6414  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 6415 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 6474  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 6475 TCTTATGCTTTTGTATACCAACCAAGAAAAATCTTTTAAAACTGGCTCTCTGATTTTAAA 6534  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 6535 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6594  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 6595 GAAAAAATTCACAAATTAACATATATTGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 6654  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 6655 ATAGAACTAGCGCCAAATTTTAAATTTTCAAAATTTTATCATACAGAAAAAATAACTAC 6714  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 6715 AAAGATATATCTCATACCTTCTAGTGAGGCTTAGGAAAAATTTTCAAAAAATATTAATTTT 6774  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 6775 AAAAAAATAAAAAATAAGAAATATTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 6834  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 6835 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 6867

RESULT 9

AAA53721	AAA53721 standard; DNA; 876 BP.	Score: 1524.00	Matches: 283
ID	AAA53721	Percent Similarity: 98.6%	Conservative: 4
XX	AAA53721	Best Local Similarity: 97.3%	Mismatches: 4
AC	AAA53721	Query Match: 97.5%	Indels: 0
XX	15-SEP-2003 (revised)	DB: 3	Gaps: 0
DT	22-DEC-2000 (first entry)		
XX	Campylobacter jejuni OH4384 CstII sialyltransferase coding sequence.	US-10-734-719-12 (1-291) x AAA53721 (1-876)	
DE			
XX	Biosynthetic locus; biosynthesis; lipid A biosynthesis;		
KW	acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;		
KW	Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;		
KW	sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;		
KW	immunity; immunogen; ganglioside; ds.		
XX			
OS	Campylobacter jejuni; OH4384.		
XX			
FH	Key Location/Qualifiers		
FT	CDS 1..876		
FT	/tag= a		
FT	/product= "CstII sialyltransferase"		
XX			
PN	MO2000046379-A1.		
XX			
PD	10-AUG-2000.		
XX			
PF	01-FEB-2000; 2000WO-CA000086.		
XX			
PR	01-FEB-1999; 99US-0118213P.		
PR	31-JAN-2000; 2000US-00495406.		
XX			
PA	(CANA ) NAT RES COUNCIL CANADA.		
XX			
PI	Gilbert M, Wakarchuk WJ;		
XX			
DR	WPI; 2000-524418/47.		
DR	P-PSDB; AAY97204.		
XX			
PT	Novel glycosyltransferase polypeptides and polynucleotides useful for		
PT	biosynthesis of ganglioside and ganglioside mimics, as diagnostic		
PT	reagents and as immunogen for producing antibodies.		
XX			
PS	Claim 6; Page 90; 120pp; English.		
XX			
CC	A reaction mixture for the synthesis of a sialylated oligosaccharide is		
CC	useful for synthesizing sialylated oligosaccharide such as ganglioside,		
CC	lysoganglioside or their mimics. Glycosyltransferases are useful for		
CC	chemo-enzymatic synthesis of oligosaccharides, including gangliosides and		
CC	other oligosaccharides that have biological activity. The enzymes and		
CC	nucleic acids that encode them are useful for studies of the pathogenesis		
CC	mechanisms of organisms that synthesize ganglioside mimics, such as C.		
CC	jejuni and the nucleic acids are used as probes to study expression of		
CC	genes involved in ganglioside mimetic synthesis. Antibodies raised		
CC	against the glycosyltransferases are also useful for analyzing the		
CC	expression patterns of these genes involved in pathogenesis. The nucleic		
CC	acids are also useful for designing antisense oligonucleotides for		
CC	inhibiting expression of the Campylobacter enzymes that are involved in		
CC	the biosynthesis of ganglioside mimics that can mask the pathogens from		
CC	the host's immune system. The oligosaccharides are useful as diagnosing		
CC	reagents or as therapeutics and as immunogens for producing antibodies.		
CC	Bacterial glycosyltransferase can be used to catalyze the formation of		
CC	oligosaccharides that are identical to the corresponding mammalian		
CC	structures and are easier and less expensive to produce in large		
CC	quantity, compared to the mammalian glycosyltransferase. The bacterial		
CC	origin of the enzymes facilitates expression of large quantities of the		
CC	enzymes using relatively inexpensive prokaryotic expression systems.		
CC	(Updated on 15-SEP-2003 to standardise OS field)		
XX			
SQ	Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	6.45e-149	Length:	876

RESULT 10  
ABT13666  
ID ABT13666 standard; DNA; 876 BP.  
XX  
AC ABT13666;

```

XX 07-FEB-2003 (first entry)
XX C. jejuni bifunctional sialtransferase cstII coding sequence #1.
XX
XX Gens; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
XX GalNAc transferase; N-Acetylgalactosamine transferase;
XX galactosyltransferase; sialyltransferase; sialic acid synthase;
XX cytidine 5'-monophosphate sialic acid synthetase;
XX CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
XX ganglioside mimetics; inflammation; tumour metastasis.
XX
XX Campylobacter jejuni.
XX
XX WO200274942-A2.
XX
XX 26-SEP-2002.
XX
XX 22-FEB-2002; 2002WO-CA000229.
XX
XX 21-MAR-2001; 2001US-00816028.
XX
XX (CAN ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
XX
XX WPI; 2003-040554/03.
XX
XX P-PSDB; ABU18479.
XX
XX New glycosyltransferases from Campylobacter, useful for synthesizing
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis
XX mechanisms of organisms that synthesize ganglioside mimetics.
XX
XX Claim 8; Page 96; 107pp; English.
XX
XX The invention comprises the amino acid and coding sequences of
XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention
XX may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
XX Acetylgalactosamine) transferase; galactosyltransferase;
XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
XX sequences of the invention are useful for ganglioside synthesis, studying
XX ganglioside mimetics, and for designing oligonucleotides to inhibit
XX expression of Campylobacter enzymes involved in the biosynthesis of
XX ganglioside mimetics that can mask the pathogen's from the host's immune
XX system. The C. jejuni oligosaccharides of the invention may be used as
XX diagnostic reagents (e.g. to locate areas of inflammation or tumour
XX metastasis). The present DNA sequence represents a Campylobacter jejuni
XX gene of the invention
XX
XX SQ Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 6,45e-149 Length: 876
XX Score: 1524.00 Matches: 283
XX Percent Similarity: 98.6% Conservative: 4
XX Best Local Similarity: 97.3% Mismatches: 4
XX Query Match: 97.5% Indels: 0
XX DB: 8 Gaps: 0
XX
XX US-10-734-719-12 (1-291) x ABT13666 (1-876)
XX
XX QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
XX |||||
XX DB 1 ATGAAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTGATTATTCAAGA 60
XX |||||
XX QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
XX |||||
XX DB 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAAATACTAT 120
XX |||||
XX QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
XX |||||
XX DB 121 CTGTGTAATAAATGCAAGCAGTATTATTACATCTCTTCTTTTGTGAACAATACTAC 180
XX |||||

```

```

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
XX |||||
DB 181 ACTTTAAACATTTTAAATCCAAATCAAGATATATGAGCCGAACATAATATGTTCTTAAT 240
XX |||||
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
XX |||||
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATATTTTCCT 300
XX |||||
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
XX |||||
DB 301 GATGCTCATTTGGGATATGATTTTTCAAACCAACTTAAAGATTTTAAATGCTTTATTTTAAA 360
XX |||||
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
XX |||||
DB 361 TTTCAGCAAAATTTATTTCAATCAAGAATTACCTCAGGGGTTTATATGTCGAGTAGCC 420
XX |||||
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
XX |||||
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAAATGGGTCA 480
XX |||||
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
XX |||||
DB 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAATTTGGCTCCTAAATTTTAAAAAT 540
XX |||||
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
XX |||||
DB 541 GATAATTCACACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
XX |||||
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
XX |||||
DB 601 GAAAAAACTTCAAAAATAAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660
XX |||||
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
XX |||||
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTCAATTTTATCATACAAGAAAAATACTACACT 720
XX |||||
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
XX |||||
DB 721 AAAGATATATCTATACCTCTAGTGAGGCTTATGGAATTTTCAAAAATATATTAATTT 780
XX |||||
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
XX |||||
DB 781 AAAAAAATAAAAAATTAAGAAAAATATTATTACAGTTGATAAAAGATCTATTAAAGATTA 840
XX |||||
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
XX |||||
DB 841 CCTGTGATATAAAGCATTATTTCAAAGGAAAA 873
XX |||||
XX RESULT 11
XX AAAS3720
XX ID AAAS3720 standard; DNA; 11474 BP.
XX
XX AC AAAS3720;
XX
XX DT 15-SEP-2003 (revised)
XX DT 22-DEC-2000 (first entry)
XX
XX DE LPS core biosynthesis locus.
XX
XX KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
XX acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
XX Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
XX sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
XX immunity; immunogen; ganglioside; ds.
XX
XX OS Campylobacter jejuni; OH4384.
XX
XX FH Location/Qualifiers
XX CDS 350..1237
XX FT /*tag= a
XX FT /product= "acyltransferase"
XX FT /note= "Open reading frame 2b"

```



```
FT CDS 1234. .2490
FT /*tag= b
FT /product= "glycosyl transferase"
FT /note= "Open reading frame 3a"
FT CDS 2786. .3955
FT /*tag= c
FT /product= "Glycosyl transferase"
FT /note= "Open reading frame 4b"
FT CDS 4025. .5068
FT /*tag= d
FT /product= "Beta 1,4-GalNAc transferase"
FT /note= "Open reading frame 5a"
FT CDS 5054. .5959
FT /*tag= e
FT /product= "Beta 1,2- galactosyltransferase"
FT /note= "Open reading frame 6a"
FT CDS 6048. .6923
FT /*tag= f
FT /product= "CstII sialyltransferase"
FT /note= "Open reading frame 7a"
FT CDS 6924. .7964
FT /*tag= g
FT /product= "Sialic acid synthase"
FT /note= "Open reading frame 8a"
FT CDS 8020. .9079
FT /*tag= h
FT /product= "Sialic acid biosynthetic enzyme"
FT /note= "Open reading frame 9a"
FT CDS 9076. .9741
FT /*tag= i
FT /product= "CMP-sialic acid synthetase"
FT /note= "Open reading frame 10a"
FT CDS complement(10554. .11366)
FT /*tag= j
FT /product= "Lipooligosaccharide biosynthetic enzyme"
FT /note= "Open reading frame 12a"
XX WO2000046379-A1.
XX 10-AUG-2000.
XX 01-FEB-2000; 2000WO-CA000086.
XX 01-FEB-1999; 99US-0118213P.
XX 31-JAN-2000; 2000US-00495406.
XX (CANA ) NAT RES COUNCIL CANADA.
XX Gilbert M, Wakarchuk WW;
XX WPI; 2000-524418/47.
XX P-PSDB; AAY97200, AAY97201, AAY97202, AAY97203, AAY97204, AAY97205,
XX AAY97206.
XX Novel glycosyltransferase polypeptides and polynucleotides useful for
XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic
XX reagents and as immunogen for producing antibodies.
XX Claim 1; Page 86-90; 120pp; English.
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is
XX useful for synthesising sialylated oligosaccharide such as ganglioside,
XX lyso-ganglioside or their mimics. Glycosyltransferases are useful for
XX chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
XX other oligosaccharides that have biological activity. The enzymes and
XX nucleic acids that encode them are useful for studies of the pathogenesis
XX mechanisms of organisms that synthesize ganglioside mimics, such as C.
XX jejuni and the nucleic acids are used as probes to study expression of
XX genes involved in ganglioside mimetic synthesis. Antibodies raised
XX against the glycosyltransferases are also useful for analyzing the
XX expression patterns of these genes involved in pathogenesis. The nucleic
XX acids are also useful for designing antisense oligonucleotides for
XX inhibiting expression of the Campylobacter enzymes that are involved in
```

CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1-7e-147	Length:	11474
Score:	1524.00	Matches:	283
Percent Similarity:	98.6%	Conservative:	4
Best Local Similarity:	97.3%	Mismatches:	4
Query Match:	97.5%	Indels:	0
DB:	3	Gaps:	0

US-10-734-719-12 (1-291) x AA53720 (1-11474)

Qy 1 MethyllysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 6048 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6107

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 6108 CTACCAATGATTGATGATTTAGATGTAATCAATTTTATTATTGGAAGATAATACATAT 6167

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
Db 6168 CTTGGTAAAAAATGCAAGGCAGTATTTTACAATCCTATTCTTTTGTGAACAATACTAC 6227

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
Db 6228 ACITTAACATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTGTTCTAAT 6287

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 6347

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 6348 GATGCTCATTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAA 6407

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 6408 TTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 6467

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGluAsnGlySer 160  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 6528 TCTTATGCTTTTGATCTAAACAAAAATCTTTTAAAAATTTGGCTCTCTTAAATTTAAAAAT 6587

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 6588 GATTAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6647

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 6648 GAAAAAATTTACAAAAATAAACTATATTGCTTATGTCCTCAACAGTCCTTTTACGAAATTTT 6707

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACTACT 6767

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

6768 AAAGATATATCATCATCTTCTAGTGAGGCTTATGCAAAATTTTCAAAAAATATTAATTTT 6827  
261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
6828 AAAAAAATAAAATTAAGAAATATTTTATTAACAGTTGATAAAGATCTATTAAGATTA 6887  
281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
6888 CCTAGTATATAAGCAATTTTCAAGGAAAA 6920

RESULT 12  
ABT13665  
ID ABT13665 standard; DNA; 11474 BP.  
XX ABT13665;  
AC ABT13665;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni genomic lipooligosaccharide biosynthesis locus.  
XX  
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylglucosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.  
XX  
XX WO200274942-A2.  
XX  
XX 26-SEP-2002.  
XX  
XX 22-FEB-2002; 2002WO-CA000229.  
XX  
XX 21-MAR-2001; 2001US-00816028.  
XX  
XX (CANVA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WJ;  
XX  
XX WPI; 2003-040554/03.  
XX  
XX  
XX New glycosyltransferases from Campylobacter, useful for synthesizing  
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
XX PT mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
XX Claim 1; Page 89-95; 107pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of  
XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
XX may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
XX Acetylglucosamine) transferase; galactosyltransferase;  
XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
XX sequences of the invention are useful for ganglioside synthesis, studying  
XX ganglioside mimetics, and for designing oligonucleotides to inhibit  
XX expression of Campylobacter enzymes involved in the biosynthesis of  
XX ganglioside mimetics that can mask the pathogen's from the host's immune  
XX system. The C. jejuni oligosaccharides of the invention may be used as  
XX diagnostic reagents (e.g. to locate areas of inflammation or tumour  
XX metastasis). the present DNA sequence represents a Campylobacter jejuni  
XX gene of the invention  
XX  
XX Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1.7e-147	Length:	11474
Score:	1524.00	Matches:	283
Percent Similarity:	98.6%	Conservative:	4
Best Local Similarity:	97.3%	Mismatches:	4
Query Match:	97.5%	Indels:	0
DB:	8	Gaps:	0

US-734-719-12 (1-291) x ABT13665 (1-11474)  
QY 1 MetLysLysValIleIlealaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGA 6107  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 6108 CTACCAATGATTTTGATGCTATTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 6167  
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
Db 6168 CTTGGTAAAAAATGCAAGGCAGTATTATTACAATCTCTATTCTTTTGAACAATACTAC 6227  
QY 61 ThrLeuLysHisIleIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 6228 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAATAATTATGTTCTTAAT 6287  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100  
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAACAACTTTTACGATTATTTTCTT 6347  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 6348 GATGCTCATTTGGGATATGATTTTTCACAAACACTTAAAGATTTTAAATGCTTATTTAAA 6407  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 6408 TTTTCACGAAATTTATTTCAATCAAGAAATATGAGACCGAATAATTATGTTCTTAAT 6467  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGTCA 6527  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 6528 TCTTATGCTTTTGATACTAAACAAAAAATCTTTTAAAAATGGCTCTCTTAATTTAAAAAT 6587  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 6588 GATAATTCACACTATATCGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 6647  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 6648 GAAAAAATTCACAAATATAAATATATGCTTATGTCCTTAACAGCTCTTTTACGAAATTTT 6707  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 6708 ATAGAACTAGCGCCAAATTTTAAATTCAAATTTTATCATACAGAAAAATAACTACACT 6767  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 6768 AAAGATATATCATCATCTTCTAGTGAGGCTTATGCAAAATTTTCAAAAAATATTAATTTT 6827  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 6828 AAAAAAATAAAATTAAGAAATATTTTATTAACAGTTGATAAAGATCTATTAAGATTA 6887  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 6888 CCTAGTATATAAGCAATTTTCAAGGAAAA 6920

## RESULT 13

AAA53725  
ID AAA53725 standard; DNA; 876 BP.  
XX AAA53725;  
AC AAA53725;  
XX  
DT 15-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
XX Campylobacter jejuni O:41 serotype alpha-2,3-sialyltransferase.  
DE

XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
 KW immunity; immunogen; ganglioside; ds.  
 XX Campylobacter jejuni; O:41 serotype.  
 OS  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..876  
 FT /\*tag= a  
 FT /product= "alpha-2,3-sialyltransferase"  
 XX  
 XX WO200046379-A1.  
 XX  
 XX 10-AUG-2000.  
 XX  
 XX 01-FEB-2000; 2000WO-CA000086.  
 XX  
 XX 01-FEB-1999; 99US-0118213P.  
 PR 31-JAN-2000; 2000US-00495406.  
 XX  
 XX (CAN ) NAT RES COUNCIL CANADA.  
 PA Gilbert M, Wakarchuk WW;  
 XX  
 XX WPI; 2000-524418/47.  
 DR P-PSDB; AAY97211.  
 XX  
 XX Novel glycosyltransferase polypeptides and polymucleotides useful for  
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 PT reagents and as immunogen for producing antibodies.  
 XX  
 XX Claim 6; Page 94; 120pp; English.  
 XX  
 XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
 CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyse the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 15-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 9.01e-148 Length: 876  
 Score: 1513.00 Matches: 279  
 Percent Similarity: 98.3% Conservatives: 7  
 Best Local Similarity: 95.9% Mismatches: 5  
 Query Match: 96.8% Indels: 0  
 DB: 3 Gaps: 0  
 US-10-734-719-12 (1-291) x AAA53725 (1-876)  
 QY 1 MeLysLysVallelleAlaGlyAsnGlyProSerLeuLysGluLleAspTyrSerArg 20

1 ATCAAAAGAGTATTATTGCTGGAATGGACCAAGATTTTAAAGAAATGATTATTCAAGA 60  
 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
 61 CTACCAATGATTGATGATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 41 LeuGlyLysCysCysLysAlaValPheTyrThrProGlyPhePhePheGluGlnTyrTyr 60  
 121 CTTGGTAAAAATGCAAAAGCAGTATTTTACAATCCTAGTCTTTTTTTTGAACAATAC 180  
 61 ThrLeuLysHisLeuLleGlnAsnGlnGluTyrGluThrGluLeuLleMetCysSerAsn 80  
 181 ACTTTAAACATTAAATCCAAAATCAAGAAATATGAGACCGAACTAATCATGTGTTCTAT 240  
 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 241 TTTAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAAACTTTTACGATTATTTCT 300  
 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 301 GATGCTCATTTGGGATATGATTTTCAACAACTTAAAGAAATCAATGCTTATTTTAAA 360  
 121 PheHisGluLleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 361 TTTCAACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTGCACAGTAGCC 420  
 141 IleAlaLeuGlyTyrLysGluLleTyrLeuSerGlyLleAspPheTyrGlnAsnGlySer 160  
 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGATCA 480  
 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
 481 TCTTATGCTTTTGATACCAACAAACAAACAAACCTTTTAAATTTGGCTCTCTTAATTTAAAAT 540  
 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLleLysAlaLeuGluPheLeu 200  
 541 GATAATTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600  
 201 GluLysThrTyrLysLleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 601 GAAAAAATTTACGAAATAAAGCTATATTGTTTATGTCCTAACAGTCTTTTACGAAATTTT 660  
 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATAAGAAAAAATACTATATCT 720  
 241 LysAspLleLeuLleProSerSerGluAlaTyrGlyLysPheSerLysAsnLleAsnPhe 260  
 721 AAAGATATATCTATCTCTCTAGTGGCTTATGGAAAAATTTACAAAAAATATTATTTT 780  
 261 LysLysLleLysLysGluAsnValTyrTyrLysLeuLleLysAspLeuLeuArgLeu 280  
 781 AAAAAAATAAAAAATTAAGAAAAATATTATTACAAAGTTGATAAAAAAGATCTATTAAAGATTA 840  
 281 ProSerAspLysHisTyrPheLysGlyLys 291  
 841 CCTAGTATATAAGCATTATTTCAAAAGGAAAA 873

RESULT 14  
 ABT13668  
 ID ABT13668 standard; DNA; 876 BP.  
 XX  
 AC ABT13668;  
 XX  
 DX 07-FEB-2003 (first entry)  
 XX  
 DX C. jejuni bifunctional sialtransferase cstII coding sequence #3.  
 XX  
 KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylglactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;

KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 XX  
 PN WO200274942-A2.  
 XX  
 XX 26-SEP-2002.  
 PD  
 XX 22-FEB-2002; 2002WO-CA000229.  
 PF  
 XX 21-MAR-2001; 2001US-00816028.  
 PR  
 XX (CAN ) NAT RES COUNCIL CANADA.  
 XX  
 XX Gilbert M, Wakarchuk WW;  
 PI  
 XX WPI; 2003-040554/03.  
 DR P-PSDB; ABJ18481.  
 XX  
 XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 XX Claim 8; Page 97; 107pp; English.  
 PS  
 XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 CC Acetylglactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention  
 XX  
 SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9,01e-148 Length: 876  
 Score: 1513.00 Matches: 279  
 Percent Similarity: 98.3% Conservative: 7  
 Best Local Similarity: 95.9% Mismatches: 5  
 Query Match: 96.8% Indels: 0  
 DB: 8 Gaps: 0

US-10-734-719-12 (1-291) x ABT13668 (1-876)

QY	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGTTATTATTCTGTAATGACCAAGTTTAAAGAAATTCATTATTCAAGA	60
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAAAATGATTTCATGATTAATGATGCAATCAATTTTATTTTGAAGATAAATACTAT	120
QY	41	LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr	60
Db	121	CTTGGTAAAAAATGCAAGACGATTTTACAAATCCTAGTCTTTTTCGAAACAATACTAC	180
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAACTAATCATGTGTTCTAAT	240
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TTTACCAGCTCACTACAGAAATCAAAATTTTGTAAGAACTTTTACGATTTATTTCTCT	300

QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTCAAACAACATTAAAGAAATTCATGCTTATTTTAAA	360
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTATTTCATCAATCAAGAAATTTACCTCAGGGCTTATATGTGTCACAGTAGCC	420
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAAATGGATCA	480
QY	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCCTATGCTTTTGTATACCAACCAAAAAATCTTTTAAAAATGGCTCCTCTAATTTTAAAAAT	540
QY	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA	600
QY	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAAAAAATTTACGAAATATAAGCTATATGTGTTATGCTCTAACAGTCTTTTAGCAAAATTTT	660
QY	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCAAAATTTTAAATTTCAATTTTATCATCAAGAAAAAATACTACTACT	720
QY	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATATCTCATCATCTCTAGTGGCTTATGGAAAAATTTACAAAAAATATTAAATTTT	780
QY	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu	280
Db	781	AAAAAAATAAAAATTAAGAAAAATATTATTACAAAGTTGATAAAGATCTATTAAAGATTA	840
QY	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTATATAAAGCAATTAATTTCAAGGAAAA	873
RESULT 15			
ID	AAZ25693		
XX	AAZ25693 standard; DNA; 1293 BP.		
AC	AAZ25693;		
XX			
DT	05-JAN-2000 (first entry)		
XX			
DE	Campylobacter jejuni alpha-2,3-sialyltransferase gene.		
XX			
KW	Campylobacter jejuni; alpha-2,3-sialyltransferase; cst-I; acceptor;		
KW	lipopolysaccharide; galactose residue; sialic acid molecule; ss.		
XX			
OS	Campylobacter jejuni.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1293	
FT		/*tag= a	
XX	WO9949051-A1.		
XX			
PD	30-SEP-1999.		
XX			
PF	22-MAR-1999; 99WO-CA000238.		
XX			
PR	20-MAR-1998; 98US-0078891P.		
PR	18-MAR-1999; 99US-00272960.		
XX			
PA	(CAN ) NAT RES COUNCIL CANADA.		
XX			
PI	Gilbert M, Wakarchuk WW;		
XX			
DR	WPI; 1999-601216/51.		

DR P-PSDB; AAY45221.

XX New sialyltransferases useful for adding sialyl residues to acceptor  
PT molecules.

XX Claim 7; Fig 2; 47pp; English.

XX The present sequence represents the Campylobacter jejuni cst-I gene which  
CC encodes alpha-2,3-sialyltransferase. The alpha-2,3- sialyltransferase  
CC protein is useful for producing desired carbohydrate structures by  
CC contacting the acceptor molecule (which has a terminal galactose residue)  
CC with an activated sialic acid molecule. The terminal galactose residue is  
CC linked to a second residue (Glc or a GlcNAc, or GlcNAc or GalNAc) in the  
CC acceptor molecule through a beta-1,3 or beta-1,4 linkage, respectively.  
CC The activated sialic acid is CMP-Neu5Ac. The polynucleotides and  
CC polypeptides facilitate the improved production of desired structures and  
XX nucleic acids encoding sialyltransferases

SQ Sequence 1293 BP; 515 A; 151 C; 176 G; 451 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	4.85e-68	Length:	1293
Score:	749.00	Matches:	148
Percent Similarity:	66.3%	Conservative:	45
Best Local Similarity:	50.9%	Mismatches:	80
Query Match:	47.9%	Indels:	18
DB:	2	Gaps:	5

US-10-734-719-12 (1-291) x AA225693 (1-1293)

QY	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
DB	46	ATGCAAAATATATCATAGCAGGAATGGACCTAGCCATAAAATATTAATATAAAGA	105
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
DB	106	CTGCCTAGAGAATATGATGTTTTTAGTGTAACAGTTTTTATTTGAAGATAAGTATTAT	165
QY	41	LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr	60
DB	166	TTAGGAAAAAAGATTAAAGCAGTATTTTTTAATCCTGCTCTTTTACACAGTATCAC	225
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
DB	226	ACTGCAAAACAACTTATCTACTAAATAAGATGATGAAATAAAATATTTTTTGTCTTACA	285
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
DB	286	TTTAATTTACCTTTTATTGAAGCAATGATTTTTTACATCAATTTTATATTTTTTCCCC	345
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
DB	346	GATGCAAAACCTTGGCTATGAAGTTATTGAAACCTTAAAGAAATTTTATGCTTATATAAAA	405
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
DB	406	TACAATGAAATTTATTTCATAAAGAATTTACTCGGGCGCTCTATATGTGCAATTGCT	465
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
DB	466	ATTGCATTAGGATATAAAACCACTATTATTATGGCCATTGATTTTTATGAAGCAGATGTT	525
QY	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaPro-----Asp	177
DB	526	ATTTATCCTTTTGAAGCTATGAGTACAAATATAAAACAATCTTCTCGGAATAAAAGAT	585
QY	178	PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu	197
DB	586	TTCAAAACCTTCAAATTTGT-----CATTTAAGGAATACGATATAGAAGCATTA	633
QY	198	GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu	217
DB	634	AAATTTGTTAAATCAATATACAAAGTTATATCTACCGATTGTGTGATGATTCTATTITG	693

Search completed: August 21, 2006, 15:00:34

Job time : 534 secs

QY	218	AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys---	236
DB	694	GCAAAATCATTTCCTTTTATCAATTAATTAATAACAATTTCACTTTAGAAAAAATAGCAT	753
QY	237	AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLys	256
DB	754	AATAAATCTATAAATGATATTTTATTGACTGATATAACTCTCGCGTAAGTTTTTATAAAA	813
QY	257	Asn-----IleAsnPheLysLysIle---LysileLys	266
DB	814	AATCAACTTAAGCTGATATAATAAATTTATGCTTAATTTTATAATATTTCTTCATTTCTAAA	873
QY	267	GluAsnValTyrTyrLysLeuIleLysAspLeu	277
DB	874	GATAATTTAATAAATTTTAAACAAAGAAAT	906

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 14:55:50 ; Search time 6304 seconds  
(without alignments)

3871.952 Million cell updates/sec

Title: US-10-734-719-12

Perfect score: 1563

Sequence: 1 MKKVIAGPSLKEIDYSR.....KLIKDLRLPSDIKHVFKGK 291

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp  
-Q=/abs/ABSSWEB.spool/US10734719/runat\_21082006\_141814\_12365/app\_query.fasta\_1  
-DB=EST -OPMT=fastap -SUFFIX=p2N.rst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs06p  
-USER=US10734719 @CGN 1.1.12067 @runat\_21082006\_141814\_12365 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMPOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_hic:\*  
7: gb\_est2:\*  
8: gb\_est7:\*  
9: gb\_est8:\*  
10: gb\_est9:\*  
11: gb\_gss1:\*  
12: gb\_gss2:\*  
13: gb\_gss3:\*  
14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	9.0	907	11	AZ685326
2	125	8.0	2789	6	AK165838 Mus muscu
3	122	7.8	886	13	CZ537148 SRAA-aad0
c 4	118.5	7.6	837	12	CC068914 CSU-K33r.

5	118	7.5	1013	14	DU782388
6	109.5	7.0	1020	14	DU782007
7	108	6.9	830	13	CZ543983
8	107	6.8	826	2	BM160438
c 9	107	6.8	885	11	AZ529344
10	106	6.8	556	4	CB390567
11	105.5	6.7	696	2	BM162427
12	105	6.7	820	10	DV143930
c 13	104.5	6.7	1130	12	CL024751
14	104.5	6.7	1131	14	AV411754
15	104	6.7	896	9	CX373647
16	103.5	6.6	825	2	BM159659
17	103.5	6.6	1101	14	CNS00ARU
18	103.5	6.6	1131	14	AY411753
19	103.5	6.6	1619	6	AY813556
20	103.5	6.6	2770	6	AK158192
21	103	6.6	790	8	CX227189
22	103	6.6	841	11	AZ673108
23	103	6.6	1475	12	CL023574
c 24	102.5	6.6	612	13	CL381292
25	102.5	6.6	730	8	CV987973
26	102.5	6.6	869	8	CV584612
c 27	102.5	6.6	911	8	CN762147
28	102.5	6.6	1373	1	AM090816
29	102.5	6.6	2416	6	AK044857
c 30	102	6.5	732	14	CT208557
31	102	6.5	980	8	CN762146
32	101.5	6.5	572	2	BM275565
33	101.5	6.5	656	4	CA856396
c 34	101.5	6.5	1376	10	DV791232
c 35	101	6.5	773	9	DN056173
36	101	6.5	804	11	BH670651
c 37	101	6.5	849	11	BH730136
38	101	6.5	885	8	CR574145
39	101	6.5	1101	14	CNS002PH
40	101	6.5	1242	1	AM102114
c 41	100.5	6.4	679	11	BZ395494
c 42	100.5	6.4	731	11	BZ023714
43	100.5	6.4	760	5	CD867200
44	100.5	6.4	1890	13	CL966179
c 45	100	6.4	805	14	CR174581

#### ALIGNMENTS

RESULT 1  
AZ685326  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

ENTKW33TF Entamoeba histolytica Sheared DNA  
907 bp  
linear  
GSS 14-DEC-2000  
genomic, genomic survey sequence.

AZ685326  
AZ685326.1  
GI:11822472  
GSS.  
Entamoeba histolytica  
Entamoeba histolytica  
Eukaryota, Entamoebidae; Entamoeba.  
1 (Bases 1 to 907)  
Loftus, B., Van Aken, S. and Fraser, C.  
Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library  
Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjl@fustcigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 55

High quality sequence stop: 794.			
FEATURES	Location/Qualifiers		
source	1..907		
	/organism="Entamoeba histolytica"		
	/mol_type="genomic DNA"		
	/strain="HM1:IMSS"		
	/db_xref="taxon:5759"		
	/clone_lib="Entamoeba histolytica Sheared DNA"		
	/note="Vector: pHS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barell, Oxford University Press, 1999)."		
ORIGIN			
Alignment Scores:			
Pred. No.:	5,55e-05	Length:	907
Score:	140.00	Matches:	72
Percent Similarity:	36.9%	Conservative:	48
Best Local Similarity:	22.2%	Mismatches:	93
Query Match:	9.0%	Indels:	112
DB:	11	Gaps:	15
US-10-734-719-12 (1-291) x AZ685326 (1-907)			
QY	38	LysTyrTyrLeuGlyLys-----LysCysLysAlaValPheTyrThrProGlyPhePhe	55
DB	1	CAGTTTGTGCTGGTAAAGAAAGTAATGCTCAACTCTGAACAGCTCTGTAAGCGATT	60
QY	56	PheGluGlnTyrTyrThrLeuLysHisLeuLeGlnAsnGlnGluTyrGluThrGluLeu	75
DB	61	GTAGATAGTTTACCTTTAAAGAACTA-----AAAGACTTTGAAATTTGAAATT	111
QY	76	IleMetCysSer---AsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThr	94
DB	112	AGAAATGCTAGTACTAATATATATCAAAATTTATGAAATAAAGGAATA-----	162
QY	95	PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu	114
DB	163	-----GATGCT-----TTCCTTAAATGTTTAAATAA	189
QY	115	-----PheAsnAlaTyrPheLysPheHisGluLeu-----	124
DB	190	CTTATGATGAACCCCTCAVTTTAAAGTTACAGAAATTCCTGAAGCTGAAAGAAACTATT	249
QY	125	-----TyrPheAsnGlnArgIleThrSerGly	133
DB	250	TATATTTTAAAGAGCGCTTTTAGAAACTGAAACAAATTTTAAATACGTTATTGTAGTGA	309
QY	134	ValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluLeu-----	148
DB	310	GAATATCTTTGACTATTTTAAAGCATTAATCTCTAAGTCTATCTCTACAACGACGA	369
QY	149	-----TyrLeuSerGlyLeuAspPheTyrGlnAsnGlySerSerTyr-----	162
DB	370	ATTGAATTAGTTAATGGAGTCTTACTAAATCATGTGGTAAATCTCTGAAGAAACTACA	429
QY	163	---AlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp	181
DB	430	CCACCTTATGAATAAATTTCTAAAAATTTAATTAAGCATCTACTGAGCTTAAAGTAA	489
QY	182	ArgSerHisTyrIleGly-----HisSerLysAsnThrAsp	193
DB	490	AGAAAAATAAATTAGATTGGAATGGCGTGTGTTAACTTTTGTGATACAGAAAAATATGA	549
QY	194	IleLeuAlaLeuGlu-----Phe	199
DB	550	CTAAGAAAGACCCCATGGAGTAATGCATGCATGTAAGAAATATGTAATATCTTC	609
QY	200	LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn	219
DB	610	TTGAGAATGGATATTAGACAGAAATGATTGTCTTGCTGGTGTGAAAGATTATTGTC	669
QY	220	PheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyr	239
DB	670	ATT-----AATGAAATTTATGTGTGAGACTGATGACCT	702
QY	240	ThrLysAspIleLeu-----IleProSerSerGluAlaTyrGlyLysPheSerLys	256
DB	703	ATTATGGATATTTTGTGACTGATCAATAGTGAAGATGATGACAAACAAATGAAG	762
QY	257	AsnIleAsnPheLys-----LysIleLysIleLys-----	266
DB	763	AACTTAATGAAAGGGTAATGATAAGTTAGACAGCAAAATGATTTAAATGATGATAA	822
QY	267	-----GluAsnValTyrTyr	271
DB	823	AAAGTTTCTTAATTCCTTGAAGAAATATGAAAGACTTACTATTGTAACGTATATCAT	882
QY	272	LysLeuIleLysAsp	276
DB	883	GAAATAATGAAAGAA	897
RESULT 2			
AK165838	2789 bp	linear	HTC 21-SEP-2005
LOCUS	Mus musculus lung RCB-0558 LLC cDNA, RIKEN full-length enriched		
DEFINITION	library, clone:G73005P05 product:Weakly similar to LD47309P		
	(fragment) homolog [Mus musculus], full insert sequence.		
ACCESSION	AK165838		
VERSION	AK165838.1 GI:74138998		
KEYWORDS	HTC; CAP trapper		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	1 Carninci, P. and Hayashizaki, Y.		
AUTHORS	High-efficiency full-length cDNA cloning		
TITLE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
TITLE	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
PUBMED	Normalizaton and subtraction of cap-trapper-selected cDNAs to		
REFERENCE	prepare full-length cDNA libraries for rapid discovery of new genes		
AUTHORS	Genome Res. 10 (10), 1617-1630 (2000)		
TITLE	11042159		
PUBMED	3		
REFERENCE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
AUTHORS	Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,		
TITLE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
PUBMED	Yamazaki, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
REFERENCE	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,		
AUTHORS	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,		
TITLE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
PUBMED	RIKEN integrated sequence analysis (RISA) system--384-format		
REFERENCE	sequencing pipeline with 384 multicapillary sequencer		
AUTHORS	Genome Res. 10 (11), 1757-1771 (2000)		
TITLE	11076861		
PUBMED	4		
REFERENCE	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,		
AUTHORS	Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,		
TITLE	Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,		
PUBMED	Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,		
REFERENCE	Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,		
AUTHORS	Fleischmann, W., Gaasterland, T., Gibbs, C., King, B., Kochiwa, H.,		

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schmitt, L.M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, G., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

11217851

5

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagawa, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Negashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultan, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Yang, L., Yuan, Z., Zavan, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sakai, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, B.S., Rogers, J., Birney, E. and Hayashizaki, Y.

FANTOM Consortium

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420 (6915), 563-573 (2002)

12466851

6

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Belsel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashizaki, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminec, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollas, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G.,

Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sincclair, B., Sperling, S., Stupka, E., Sugiura, K., Sultana, R., M.S., Takenaka, Y., Taki, K., Tanno, J., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plesky, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium

The transcriptional landscape of the mammalian genome

Science 309 (5740), 1559-1563 (2005)

16141072

7

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

RIKEN Genome Exploration Research Group

Antisense transcription in the mammalian transcriptome

Science 309 (5740), 1564-1566 (2005)

16141073

8 (bases 1 to 2789)

Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. 2789

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="FANTOM\_DB:G730005P05"

/db\_xref="taxon:10090"

/clone="G730005P05"

/cell\_line="RCB-0558 LLC"

/tissue\_type="lung"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

202.2178

/note="unnamed protein product; Weakly similar to LD47309P (Fragment) homolog [Mus musculus] (Uniprot|Q8BRH0, evidence: FASTY, 99.7%ID, 86.4%length, match=1953) putative"

/codon\_start=1

/protein\_id="BAE38405.1"

/db\_xref="GI:74138999"

# COMMENT

## FEATURES source

### CDS





Qy	9	AsnGlyProSerLeuLys---GluIleAspTyrSerArgLeuProAsnAspPheAspVal	27
Db	23	ATAATCCTAATGAAATAATTAATAATGATGCAATAATA-----TTTAGTTG	73
Qy	28	PheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysAla	47
Db	74	AAAAAATTAAC-----AATAAATATATCTGGT-----ACT	106
Qy	48	valPheTyrThrProGlyPhePheGlu-----	57
Db	107	ATGTTTTTATCAAGAAATTTTTTTTACAATGGTTGCCCTGTGAAGAAATATATAAG	166
Qy	58	-----GlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu	72
Db	167	AAGAATCTGGTATAAATAATATATGCTATATAACATGTTATGATAATGA-----	217
Qy	73	ThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheVal	92
Db	218	-----ATTGATCTCCAAATTAATAACATGTCACATTA-----TTTTTA	256
Qy	93	LysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeu	112
Db	257	CCTATTTCCTCAATATGAAAGAAACCTTTTTTTCATGTGCTTGGATAGAACAAACA	316
Qy	113	LysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSer	132
Db	317	TCATTACCAAAATATA-----GAAGTTGGATTT-----	343
Qy	133	GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGly	152
Db	344	-----TATCTAGAAGAAAGTTTAAAGGT	367
Qy	153	IleAspPheTyrGlnAsnGlySer-----	160
Db	368	ATTAATGAGATACTAATAATAGTTTGAATATTGAAATTAGATCAAAATTTGTATGAATTT	427
Qy	161	-----SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeu	174
Db	428	AAAAGCCAATTTATGTATTCATTCATTTATACAAATGAAATCTAATATATAATTTG	487
Qy	175	AlaProAspPheLys-----AsnAspArgSerHisTyrIleGly-----	187
Db	488	AGAGACATTTCAAAAAACATTAATAATAGTATTATTTTGGGCAAAAGTTATATTTT	547
Qy	188	-----HisSerLysAsnThrAspIleLysAlaLeuGluPheLeuLysThrTyrLysIle	206
Db	548	TTTAACATGAAATAATTTGATATTTCATATTATTTGATTTGAAGAACCT--AAAGTA	604
Qy	207	LysLeu---TyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaPro	225
Db	605	CAAAATACCATATTCAATTTGC-----	625
Qy	226	AsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIle	245
Db	626	-----ATCCAAAGAGTCTAAAGTTAT--AAAGCAATATATATTT	661
Qy	246	ProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysIle	265
Db	662	CCAGTACTGAA-----AATTTGAAGATATTAACCTT	694
Qy	266	LysGlu-----AsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAsp	283
Db	695	GACAACACAACATAATGAATCTATCAATTAATTAATCA-----TCTAAT	739
Qy	284	IleLysHisTyrPheLysGlyLys	291
Db	740	TTTAATGAACATATTAAACCAAAA	763

RESULT 4  
 CC068914 837 bp DNA linear GSS 16-APR-2003  
 LOCUS CSU-K33r.25C12.SP6 CSU-K33r Aedes aegypti genomic clone  
 DEFINITION

CSU-K33r.25C12, genomic survey sequence.  
 CC068914  
 CC068914.1 GI:29907420  
 GSS.  
 Aedes aegypti (yellow fever mosquito)  
 Aedes aegypti  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;  
 Culicidae; Culicini; Aedes; Stegomyia.  
 1 (bases 1 to 837)  
 Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.  
 End sequencing of Aedes aegypti BACS  
 Unpublished (2003)  
 Other GSSs: CSU-K33r.25C12.T7  
 Contact: Brendan Loftus  
 Department of Eukaryotic Genomics  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-3543  
 Fax: 301-838-0208  
 Email: entaetigr.org  
 Library was provided by Susan Brown and Dennis Knudson at Colorado  
 State University.  
 Seq primer: SP6  
 Class: BAC ends.  
 Location/Qualifiers  
 1..837  
 /organism="Aedes aegypti"  
 /mol\_type="genomic DNA"  
 /strain="Rexville"  
 /db\_xref="taxon:7159"  
 /clone="CSU-K33r.25C12"  
 /clone\_lib="CSU-K33r"  
 /note="Vector: pBeloBAC11; Site\_1: HindIII"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.0137 Length: 837  
 Score: 118.50 Matches: 56  
 Percent Similarity: 46.4% Conservative: 54  
 Best Local Similarity: 23.6% Mismatches: 82  
 Query Match: 7.6% Indels: 46  
 DB: 12 Gaps: 10  
 US-10-734-719-12 (1-291) x CC068914 (1-837)  
 Qy 13 LeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGln 32  
 Db 652 TTGAATTTAAAGAAATTTCTTTTCGAAACCAAGAAATTC-----TTTCGAAACCAAGA 599  
 Qy 33 PheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrPro 52  
 Db 598 ATTTCTTTTCGAAATCAAGAAATTTCTTTTCGAAA----- 566  
 Qy 53 GlyPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72  
 Db 565 -----TTCAAGAAATTTTTTTTCAAAATTCAGAAATTCGAAATTAAGAAATTTCTT 512  
 Qy 73 ThrGluLeu---IleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPhe 91  
 Db 511 TCGAAATTTAAAGAAATTTACTTTTCGAAATTCAGAAATTTTCATTCGAAA----- 464  
 Qy 92 ValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr---AspPhePheLys 110  
 Db 463 TTCAAGAAATTTTCATTCGAAATTCAGAAATTTTCATTCGAAATTCAGAAATTTCTTTCG 404  
 Qy 111 GlnLeuLysGluPheAsnAlaTyrPheLys---PheHisGluIleTyr-----PheAsn 127  
 Db 403 AAATTCAGAAATTTTCATTCGAAATTCAGAAATTTTCATTCGAAATTCAGAAATTTTCAT 344  
 Qy 128 GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGlu 147  
 Db 343 TCGAAA-----TTCAAGAA 329

```

QY 148 lIeTyr-LeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrIly 167
Db 328 TTTCAATCGAAATTCAAAGAATTTCAATCGAAATTCAAAGAATTCATTCGAAATTCAAA 269
QY 167 s-----GlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspAr 182
Db 268 GAATTCATTCGAAATTCAAAGAATTCATTCGAAATTCAAAGAATTCATTCGAAATTC 209
QY 182 gSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluY 202
Db 208 AAAGAATTCATTCGAAATTCAAAGAATTCATTCGAAATTCATTCGAAATTCATTCGAA 150
QY 202 sThrTyr-LysIleLysLeuTyrCysLysProAsnSerLeuLeuAlaAsnPheIleG 222
Db 149 ATCAATTCGAAATTCAAAGAATTCGTTGAAATTCAAAGAATTCATTCGAAATTCAAAG 90
QY 222 luLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsn 237
Db 89 AATTTCTTTAAATTTA---AAGAATTTCTTCAATATTCCAAAAT 46

RESULT 5
DU782388 1013 bp DNA linear GSS 27-JAN-2006
LOCUS ASXB4195.b2 HF500_10-06-02 uncultured marine microorganism
DEFINITION HF500_10-06-02 genomic clone HF0500_102C09, genomic survey
sequence.
ACCESSION DU782388.1 GI:85796675
VERSION DU782388.1
KEYWORDS uncultured marine microorganism HF500_10-06-02
SOURCE uncultured marine microorganism HF500_10-06-02
ORGANISM unclassified sequences; environmental samples.
REFERENCE 1 (bases 1 to 1013)
AUTHORS DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
and Karl,D.M.
TITLE Comparative genomics reveals ecological trends in stratified
microbial communities in the ocean's interior
JOURNAL Science (2006) In press
COMMENT Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong
US DOE Joint Genome Institute
US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679
Email: PMRichardson@lbl.gov; delong@mit.edu
North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002
Coordinates: 22.45 N, 158 W Depth 500 m Temperature: 7.25 C
Salinity: 34.07 psu Oxygen: 118.0 umol/kg
Class: fosmid ends
FEATURES
    source
        1..1013
            /organism="uncultured marine microorganism HF500_10-06-02"
            /mol_type="genomic DNA"
            /db_xref="taxon:361149"
            /clone="HF0500_102C09"
            /cell_type="marine picoplankton, less than 1.8 um, greater
            than 0.22 um fraction"
            /clone_lib="HF500_10-06-02"
            /notes="Vector: pCICFOS; North Pacific Subtropical Gyre
            (Hawaii) picoplankton genomic fosmid DNA library prepared
            from marine picoplankton in the less than 1.6 um, greater
            than 0.22 um fraction. Picoplankton collected at 500 m
            depth on 10/6/2002, Coordinates: 22.45 N, 158 W. Sample
            Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 500 m
            Temperature: 7.25 C Salinity: 34.07 psu Oxygen: 118.0
            umol/kg"
    ORIGIN

```

```

Alignment Scores:
Pred. No.: 0.0201 Length: 1013
Score: 118.00 Matches: 67
Percent Similarity: 38.4% Conservatives: 40
Best Local Similarity: 24.0% Mismatches: 107
Query Match: 7.5% Indels: 65
DB: 14 Gaps: 13

US-10-734-719-12 (1-291) x DU782388 (1-1013)
QY 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgIeu 21
Db 223 AAAAATCTATCTCATGTGTAGTGGAGGCCCTCGTCGGCGGACGTGTGATTGGTCAGCCCTC 282
QY 22 Pro-----AsnAspPheAspValPheArGysAsnGlnPheTyrPhe---GluAspLys 38
Db 283 CCTATTAAATCGAGATTTT---GTATGGTCATGCAATAAATCTTTTATGAATAAAGACCTT 339
QY 39 TyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGln 58
Db 340 AATACCGTCGGTGGATTTTCGCATCTATTGGCCCTACA----- 378
QY 59 TyrTyrThrIleuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCys 78
Db 379 ---GTTGATTATCATGAACTGCTCCAGGACCACATAGAAAAACATGGAACATGTTTGC 435
QY 79 SerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyr 98
Db 436 GCT-----ATTGAGGAGGATCTCTCTCTTTTGAACAAAA 471
QY 99 PheProAspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyr 118
Db 472 GAAGAGCTTATTCACCTTA-----AAAGAAATATCAAAAAATAATATGTT 516
QY 119 PheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAla 138
Db 517 TTT---TATCATCTCCGATATTTTTCG---AAACTTGAACATTTGGCGAGAATGGTTCT 570
QY 139 ValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr----- 156
Db 571 TTAGCAACTCACCTTCAAGTTAAGAAAAATTTATTTTCTGGAATGGACGGATATCCCGGA 630
QY 157 -----GlnAsnGlySer 160
Db 631 AAGCGCGGCAATTTGTATGAACATGCCTTTGAAGAGGTGCGAAGAACACCAAGGGCGA 690
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 691 ACTTTTTCGTATGATTTGCACAGAGCAATATGTTTATTTATGGGATTACTTTATGAAC 750
QY 181 AspArgSerHisTyrIle-----GlyHisSerLysAsn----- 191
Db 751 ACCTTAGGTTCCAAAGGTAAAAATATCAGATTTAGGAGAGGCCACGCGCAATCAGTCG 810
QY 192 ThrAspIleLysAlaLeu-GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLe 211
Db 811 ACTGATATTTCTAGACAAATTTATTTCTTTGGAATATACATAAAATA----- 856
QY 211 uCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPh 231
Db 857 -----TGGTATATTAACTTTAGTTTATATAGGAGAATTATGAACGATTTCATT 903
QY 231 eIlelleGlnGluLysAsn---AsnTyrThrLysAspIleLeuLeuProSerSer 248
Db 904 CATTATCGCAAGAAATCGGTATTAAATCAATATGGGGATATCGAAATAGCAAAATCT 958

RESULT 6
DU782007 1020 bp DNA linear GSS 27-JAN-2006
LOCUS ASXB4011.b2 HF500_10-06-02 uncultured marine microorganism
DEFINITION HF500_10-06-02 genomic clone HF0500_101C10, genomic survey
sequence.

```

ACCESSION	DU782007	79	SerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyr	98
VERSION	DU782007.1	436	GCT-----ATTGAAGGAGGATCTCTCTCTTTTAGAACAAAA	471
KEYWORDS	GSS	Qy		
SOURCE	uncultured marine microorganism HF500_10-06-02	Qy	PheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyr	118
ORGANISM	uncultured marine microorganism HF500_10-06-02	Db		
REFERENCE	unclassified sequences; environmental samples.	Db		
AUTHORS	1 (bases 1 to 1020)	472	GAAGAGCTTATTCACCTTA-----AAAGAAATATCAAAAAAATAATTGTT	516
	DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,	Qy		
	Frigaard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W.	Db		
	and Karl, D.W.	Qy		
TITLE	Comparative genomics reveals ecological trends in stratified	Db		
JOURNAL	microbial communities in the ocean's interior	Qy		
COMMENT	Science (2006) In press	Db		
	Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,	Qy		
	Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson	Db		
	and Edward DeLong	Qy		
	US DOE Joint Genome Institute	Db		
	2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA	Qy		
	Tel: 617-253-5271	Db		
	Fax: 617-253-2679	Qy		
	Email: PMRichardson@lbl.gov; delong@mit.edu	Db		
FEATURES	North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid	Qy		
source	DNA library prepared from marine picoplankton in the less than 1.6	Db		
	um, greater than 0.22 um fraction. Sample Date: 10/6/2002	Qy		
	Coordinates: 22.45 N, 158 W Depth 500 m Temperature: 7.25 C	Db		
	Salinity: 34.07 psu Oxygen: 118.0 umol/kg	Qy		
	Class: fosmid ends.	Db		
	Location/Qualifiers	Qy		
	1..1020	Db		
	/organism="uncultured marine microorganism HF500_10-06-02"	Qy		
	/mol_type="genomic DNA"	Db		
	/db_xref="taxon:361149"	Qy		
	/clone="HF0500_101C10"	Db		
	/cell_type="marine picoplankton, less than 1.8 um, greater	Qy		
	than 0.22 um fraction"	Db		
	/clone_lib="HF500_10-06-02"	Qy		
	/notes="vector: pC1FOS; North Pacific Subtropical Gyre	Db		
	(Hawaii) picoplankton genomic fosmid DNA library prepared	Qy		
	from marine picoplankton in the less than 1.6 um, greater	Db		
	than 0.22 um fraction. Picoplankton collected at 500 m	Qy		
	depth on 10/6/2002. Coordinates: 22.45 N, 158 W. Sample	Db		
	Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 500 m	Qy		
	Temperature: 7.25 C Salinity: 34.07 psu Oxygen: 118.0	Db		
	umol/kg"	Qy		
ORIGIN		Db		
Alignment Scores:		Qy		
Pred. No.:	0.187	Db		
Score:	109.50	Qy		
Percent Similarity:	38.4%	Db		
Best Local Similarity:	22.8%	Qy		
Query Match:	7.0%	Db		
DB:	14	Qy		
		Db		
US-10-734-719-12 (1-291) x DU782007 (1-1020)		Qy		
		Db		
Qy	2 LysLysValIleLeAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu	21		
Db		223	AAAACTATTCATGCTAGGTGGAGGCCCTCTCGCGGACGTTGATGGTCAGCCCTC	282
Qy	22 Pro-----AsnAppheAspValPheArgCysAsnGlnPheTyrPhe---	38		
Db		283	CCTATTAACTCAGATTTT---GTATGTCATGCAATAACTTTTATGAATAAAGACCTT	339
Qy	39 TyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGln	58		
Db		340	AATACCGTCGGTGTGGATTGTCATCTATTGGCCCTACA-----	378
Qy	59 TyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCys	78		
Db		379	---GTTGATTATCACATGAATGACTGCTCAGACCACCATAGAAAAACATGGNACTGTTGC	435

```

UK.
Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..830
                     /organism="Strongyloides ratti"
                     /mol_type="genomic DNA"
                     /strain="Isfemale line ED321 heterogenic"
                     /db_xref="taxon:34506"
                     /dev_stage="infective larval stage (iL3)"
                     /lab_host="GS10"
                     /clone_lib="Strongyloides ratti whole genome shotgun
                     library (SRAAGS 004)"
                     /note="Vector: pOTW13; Site 1: BstXI; Site 2: BstXI;
                     Strongyloides ratti genomic DNA was randomly sheared,
                     end-repaired and size fractionated to enrich for 2-4 kb
                     fragments. Genomic DNA was provided by Fiona Thompson
                     (F.Thompson@bristol.ac.uk) and Mark Viney
                     (Mark.Viney@bristol.ac.uk) at the University of Bristol,
                     Bristol, UK. Sequencing by Washington University Genome
                     Sequencing Center, St. Louis, MO."

ORIGIN
Alignment Scores:
Pred. No.:          0.211      Length:      830
Score:              108.00      Matches:    70
Percent Similarity: 37.2%      Conservative: 36
Best Local Similarity: 24.6%   Mismatches: 102
Query Match:        6.9%      Indels:    78
DB:                  13       Gaps:       12

US-10-734-719-12 (1-291) x CZ543983 (1-830)
QY   30  CysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCys-LysAlaValph 49
Db   104 TGTACAAAGTGGTTTACACACCAAGATTTATTTGGGAAAGTAAATAGGAAGCTGTATT 163
QY   49  eTyrThrProGlyPhePheGluGlnTyrTyr-ThrLeuLysHisLeuIleGlnAsnG 69
Db   164 TTATATGTTTGGCTTGTATGTTATTAATTTATCAGATGTGGAATCATTTAGTAAGAAC 223
QY   69  lnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnG 89
Db   224 AGCGGATCATTTAGAA-----GAAAGGG 247
QY   89  luAsnPheValLysThrPheTyrAspTyrPheProaspAlaHisLeuGlyTyrAspPhep 109
Db   248 AACGTTTGGAAAAAGCTTATTTGTGATCAAGATTAGCTGATCATCAAAACAGATGTTT 307
QY   109 heLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn--- 127
Db   308 AT-----TTTAATATTATTAACTTTTTCATGATAAAGCTAAATATAA 352
QY   128 --GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysG 147
Db   353 CTCAGAACATAAATGTAGGAA--TATATT-----G 381
QY   147 luIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrL 167
Db   382 AAAGATATATA----- 392
QY   167 ysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleG 187
Db   393 -----AATTTATCCAAACAAATTCATGATATTTTGGATGATCTCGAAGGAAGTAG 444
QY   187 lyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGlu-----LysThrTyrL 205
Db   445 ATTTTGAAGAAGATAGAAATGATGAAGTCACATCCATTTGGGAATTTTATAAAGATTACA 504
QY   205 ys----lleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuA 224
Db   505 AAGATTATAAAGATTAT-----TACAAAGTTAATACAGATGAAGTTG 546
QY   224 laProAsnLeuAsnSerAsnPhe-----llelleGlnGluLysAsnAsnTyrThrLysA 242

```

```

Db   547 GACCCCTCAATTACACAGTGAATTTTCAAGAGGTGTATGAAGAAATTTGGAATGAAGCTGACA 606
QY   242 spile-LeulleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe--- 260
Db   607 TGATAGATTTTACAGATGAGGAAGGATATGGAATAATTTTAGATTTTAAACAAATTTACT 666
QY   261 -----LysLysIleLysIleLysGluAsnValTyrTyrLysLeu 273
Db   667 ACTATCCCTATTTTAAATTTAAAGATGTTAAGGTAAAGATAATACTTTTTTTACATA 726
QY   274 Ile-----LysAsp-----LeuLeuArgIleProSerAspIleLysHis 286
Db   727 ATTTTATTTATTTGTAGAAAAATTTGAATTTATCTTGAAGTATGTTTCCAGTTTGAATAT 786
QY   287 TyrPheLys 289
Db   787 TTTTGGAAA 795

RESULT 8
BM160438
LOCUS      826 bp      mRNA      linear      EST 04-DEC-2001
DEFINITION EST562961 PyBS Plasmodium yoelii yoelii cDNA clone PJCR30 5' end,
            mRNA sequence.
ACCESSION  BM160438
VERSION    BM160438.1 GI:17306119
KEYWORDS   EST.
SOURCE     Plasmodium yoelii yoelii
ORGANISM   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE  1 (bases 1 to 826)
AUTHORS   Carlton,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Vaidya,A.B.,
            Fraser,C.M. and Carucci,D.J
TITLE     Plasmodium yoelii EST project at TIGR
JOURNAL   Unpublished (2001)
COMMENT   Contact: Jane Carlton
            Parasite Genomics Group
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-530-9319
            Fax: 301-838-0208
            Email: carlton@tigr.org
            For clone info, please contact the Malaria Research and Reference
            Reagent Resource Center, ATCC
            http://www.malaria.mr4.org/mr4pages/index.html
            Seq primer: ADF.
FEATURES             Location/Qualifiers
     source           1..826
                     /organism="Plasmodium yoelii yoelii"
                     /mol_type="mRNA"
                     /strain="17XL"
                     /sub_species="yoelii"
                     /db_xref="taxon:73239"
                     /clone="PYCJR30"
                     /dev_stage="Asexual blood stages"
                     /lab_host="E. coli XL-1 Blue"
                     /clone_lib="PyBS"
                     /notes="Vector: PAD-GAL4; At 20-25% parasitemia, blood was
                     collected from BALB/cBYJ mice infected with Py17XL
                     parasites, and leukocytes removed by passage over
                     microcrystalline cellulose columns. Total RNA was
                     isolated using the guanidinium isothiocyanate method, and
                     mRNA isolated using oligo(dT)-cellulose chromatography.
                     First strand cDNA synthesis was completed using a 50-base
                     primer and reverse transcriptase in the presence of
                     5-methyl dCTP. After second strand synthesis, uneven
                     termini were treated with pfu DNA polymerase and EcoRI
                     adaptors ligated to the blunt ends. The sample was cleaved
                     with XhoI and separated on a Sephacryl S-500 column.
                     Size-fractionated cDNA was precipitated and ligated to
                     HybridZAP arms directionally using EcoRI-XhoI cleaved arms.
                     After packaging, the phagemid vector (pAD-GAL4) was
                     excised from the HybridZAP vector and plasmid DNA

```

isolated."	
ORIGIN	
Alignment Scores:	
Pred. No.:	0.272
Score:	107.00
Percent Similarity:	35.1%
Best Local Similarity:	22.3%
Query Match:	6.8%
DB:	2
US-10-734-719-12 (1-291) x BM160438 (1-826)	
Qy	62 LeuLysHisLeuIleGlnAsnGlnGlyThrGluThrGluLeuLeuMetCysSerAsnTyr 81
Db	91 TTAGAGCATCTACCCAAATGAAGAA-----ACCTCAAAAATAT 132
Qy	82 -----AsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhe 99
Db	133 ATGAACAATTTGTTTTTATCAAAAATAATGAATTAATAAAAAATAGACTCATCATATTTT 192
Qy	100 ProAspAlaHisLeuGly-----TyrAspPhe 108
Db	193 CAAGATAAATATTAGGATATTATTTTGGAGCCTCATGTGTAATAATTTGTGTATCATTT 252
Qy	109 PheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGln 128
Db	253 ATAAATAATTTAAATTTATTTAAACACTATTTCCTTTTGTGAAATTT----- 300
Qy	129 ArgIleThrSerGlyValTyrMetCysAlaValAlaLeuLeuGlyTyrLysGluLeu 148
Db	301 -----ATATATATACCATTTGTATCAACA 324
Qy	149 TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAla-----PheAspThr 166
Db	325 TATACAGATTATTAATTTTAAAAAATACAAATTTTATATCTTACCTTTTGAT--- 381
Qy	167 LysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIle 186
Db	382 -----AATTATTTATATAGCTAATAAATTTAA----- 411
Qy	187 GlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIle 206
Db	412 -----GTCAAAAAT----- 420
Qy	207 LysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsn 226
Db	421 -----TTGCCATCTTTTATTTATATAGCACCA--- 447
Qy	227 LeuAsnSerAsnPheIleIleGlnGlu-----LysAsnAsnTyrThr 240
Db	448 ---AATAATAATATCTTGTAGGAGTGGAGTACAAATTAATAAACTGCAATTTATGTA 504
Qy	241 LysAspIle-----LeuIleProSerSerGluAlaTyrGlyLys---PheSerLysAsn 257
Db	505 AATAAATTTCAAACTCTTTGATAAAAAATATACAATACACCCCAAAACATTCAAATCAAAT 564
Qy	258 IleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuLysAspLeu 277
Db	565 AATCGATT-----TTCGACTTATTTACAAATTAGTTATCTTCCATC 606
Qy	278 LeuArgLeuProSerAspIleLysHisTyrPhe 288
Db	607 CTATATATCTCAACAAGTCTTAATATTTTTTT 639
RESULT 9	
AZ529344/c	
LOCUS	885 bp DNA linear GSS 03-NOV-2000
DEFINITION	ENTB239TR Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, genomic survey sequence.
ACCESSION	AZ529344
VERSION	AZ529344.1 GI:11082116
KEYWORDS	GSS.
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
Entamoeba histolytica	
Entamoeba histolytica	
Eukaryota; Entamoebidae; Entamoeba.	
1 (bases 1 to 885)	
Loftus,B., Van Aken,S. and Fraser,C.	
Determination of clone end sequences from Entamoeba histolytica	
HM1:IMSS sheared DNA library	
Unpublished (2000)	
Contact: Brendan J Loftus	
Department of Eukaryotic Genomics	
The Institute for Genomic Research	
9712 Medical Center Dr., Rockville, MD 20850, USA	
Tel: 301 838 0208	
Fax: 301 838 3543	
Email: b.loftus@tigr.org	
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared	
DNA library	
Seq primer: M13-Reverse	
Class: shotgun	
High quality sequence start: 17	
High quality sequence stop: 811.	
FEATURES	
source	
1..885	
/organism="Entamoeba histolytica"	
/mol_type="genomic DNA"	
/strain="HM1:IMSS"	
/db_xref="taxon:5759"	
/clone_lib="Entamoeba histolytica Sheared DNA"	
/note="Vector: pHS01; Site 1: Bst I; Constructed at The	
Institute for Genomic Research (TIGR), Rockville, MD.	
Genomic DNA isolated from broth cultures of E. histolytica	
using a method described by Clark and Diamond (Clark,	
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a	
method for isolate identification. Exp. Parasitol.	
77:450.). The DNA was mechanically sheared to give a	
tight size distribution (~2 kb). The v + i method used for	
the library construction is described in detail in Smith,	
H.O. and Venter, J.C. (Making small insert libraries for	
whole genome shotgun sequencing projects. In Genome	
Sequencing: A Practical Approach, eds. M. Vaudin and B.	
Bartell, Oxford University Press, 1999)."	
ORIGIN	
Alignment Scores:	
Pred. No.:	0.298
Score:	107.00
Percent Similarity:	34.4%
Best Local Similarity:	23.5%
Query Match:	6.8%
DB:	11
US-10-734-719-12 (1-291) x AZ529344 (1-885)	
Qy	21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db	679 TTGCCATCTCTTTTCTGTTTATAATTTCAATACACTTATATGACAGTGTCCATAT 620
Qy	41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db	619 TATGGT-----ATTTTA 608
Qy	61 ThrLeuLysHisLeuIleGlnAsnGlnGlu----- 70
Db	607 ACAATATTCATCTCTTAACACACGCAAAATTTGTAATTTCTTACTGTGACTCAACAGT 548
Qy	71 -----TyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeu 86
Db	547 TCAATTGCAACATATATTTATGAAATTTATTTGTTGGAACAATTAATGAATATATAT 488
Qy	87 GluAsnGluAsnPhe-----ValLysThrPheTyrAspTyrPheProAspAla 102
Db	487 AAATCTCGTTCAAGTGGTGGGTGGCATCTATTCAAGATATAGTTTTTAGTGTTCGAACA 428

QY	103	HisLeuGlyTyrAspPhePheLeuGlnLeuLysGluPheAsnAlaTyrPheLysPheHis	122
Db	427	GATAAAATATTATTTTATTTTATTTTATTTTCAAAA-----GTTCAATCTCTTTTCA	377
QY	123	GluLeuPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAla	142
Db	376	TTATATTATTTTCAATCGTTGTTATTTTGGTAAGTTTAAATGTTTGTATTATAAG	317
QY	143	----LeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer	161
Db	316	TATTATCATATAAAGAAGAAGATTTAACAAATATACCTTATTTTCATATTCAATCTCTC	257
QY	162	TyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp	181
Db	256	CCATTTTATCATGAAAAATGAATAAAT-----TCTATTAAACCCCTCTCTTTAACGAAGAT	203
QY	182	ArgSerHis-----TyrIleGlyHisSerLysAsnThrAsp	193
Db	202	TGTATTTCATGATGAGTATGCTATTGGTATACATTTAAAGGAGATGAAATCAATCAAC	143
QY	194	IleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysPro	213
Db	142	ATATCT-----GATCTTTAATTAAC-AATPAATGAAGCTACTGAAC	78
QY	214	AsnSerLeuLeuAlaAsnPhelIleGluLeuAlaProAsnLeuAsnSerAsnPhelIle	233
Db	130	-----AACTCTTTTAGTAATTTT-----	113
QY	234	GlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr---Gly	252
Db	112	-----GATCTTTAATTAAC-AATPAATGAAGCTACTGAAC	78
QY	253	LysPheSerLysAsnIleAsnPhelLysLysIleLysLeuGluAsnValTyrTyrLys	272
Db	77	AAATGCGAGT-----TATATCAAAA-AATGTTTATTACAAG	43
QY	273	LeuIleLysAspLeu	277
Db	42	ATTATAATGCGGTTG	28
RESULT 10			
CB390567			
LOCUS			
DEFINITION			
CB390567_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.			
ACCESSION			
CB390567.1 GI:30732277			
VERSION			
EST.			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
Genome annotation and resource 1.1: experimental verification of the expression			
JOURNAL			
COMMENT			
Nat. Genet. (2003) In press			
Contact: Vidal M			
Marc Vidal Laboratory			
Dana Farber Cancer Institute			
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA			
Tel: 617 632 5180			
Fax: 617 632 5739			
Email: Marc.Vidal@dfci.harvard.edu			
Sequence tag of Gateway entry clones. The primers used were			
designed on the predicted protein encoding ORF C. elegans ORFeome			
Cloning project : Contact david.hill@dfci.harvard.edu or			
marc_vidal@dfci.harvard.edu			

FEATURES		POLYA=No.		Location/Qualifiers	
source		1..556		/organism="Caenorhabditis elegans"	
				/mol_type="mRNA"	
				/strain="N2"	
				/db_xref="taxon:6239"	
				/sex="Hermaphrodite and male"	
				/tissue_type="whole animal"	
				/dev_stage="mixed stage"	
				/clone_lib="AD-wrmcDNA"	
				/notes="The AD-wrmcDNA library was generated with poly(A) + RNA isolated from both hermaphrodite and male N2 worms of all larval stages, embryos, adults and dauers and the subsequent generation of cDNAs by poly(A) priming. The cDNAs were cloned into pPC86"	
ORIGIN		US-10-734-719-12 (1-291) x CB390567 (1-556)		Alignment Scores:	
		Pred. No.:		Length:	
		Score:		Matches:	
		Percent Similarity:		Conservative:	
		Best Local Similarity:		Mismatches:	
		Query Match:		Indels:	
		DB:		Gaps:	
QY		7		AlaGlyAsnGlyProSerLeuLysGluIleAsp---TyrSerArgLeuProAsnAspPhe	
Db		20		AGTGGTGAAGTGGCGACATTTGATGACTTGAACAATTCACGGAATGTTGAATCCCTAT	
QY		26		AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLys---	
Db		80		-----TCACAACCT-----TCAATAAATTTTATTACGGAAGAAAGAA	
QY		45		-----CysLysAlaValPheTyrThrProGlyPhePhe-----Phe	
Db		119		AGGTATATCCGAAGCGTTTGGTAGTAATTCGGCATATTTTCAATATGTTTCAGTATT	
QY		57		GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGlyTyrGluThrGluLeuIle	
Db		179		GATAAATATCAATATTAATGAAGAAATTTGAATCGAAGAAATGTAATGATGTACTGCA	
QY		77		MetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyr	
Db		239		GAATGCAACAATATACCT-----TTTAATCAAAACATTTGTTCCACCATTTTAT	
QY		97		AspTyrPheProAspAlaHisLeuGly-----TyrAspPhePheLysGlnLeuLysGlu	
Db		287		GATTATCTTCAAGATTTTACAATTTCTCCTCGTTAIGAT-----	
QY		115		PheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal	
Db		326		-----ATA	
QY		135		TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp	
Db		329		TCTTTTGTCTTATTCCAAAAGTTGTG-----TCTACAATCGGAACGCGCAATATGT	
QY		155		PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeu	
Db		383		TACATACAAGATCCGGAGGCTTTTCACC-----	
QY		175		AlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIle	
Db		410		-----AAAAACAACAGA-----	
QY		195		LysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyr-----CysLeuCys	
Db		422		-----ACAATATCAACGGAATGATGTGTGGGAAGATTTTTC	
QY		213		ProAsnSerLeuLeuAlaAsnPhelIleGluLeuAlaProAsnLeuAsnSerAsnPhe---	

```

Db      458 GAAAAAATGAATGAAATCTTTTAAATGTCGACGAGTTTGAATCTTAACTTTGAT 517
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      232 ---lilelclnGlulysAsnTyrThrLysAspTle 243
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      518 AATATAATTTTCACAGAAATCCCTATGATCGGTTCATT 556
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::|||

RESULT 11
LOCUS   BM162427
DEFINITION EST564950 PyBS Plasmodium yoelii yoelii cDNA clone PYCKS34 5' end,
mRNA sequence.
ACCESSION BM162427
VERSION   1
KEYWORDS  EST.
SOURCE    Plasmodium yoelii yoelii
ORGANISM  Plasmodium yoelii yoelii
REFERENCE 1 (bases 1 to 696)
AUTHORS   Carlson,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Vaidya,A.B.,
Fraser,C.M. and Carucci,D.J.
TITLE     Plasmodium yoelii EST project at TIGR
JOURNAL   Unpublished (2001)
COMMENT   Parasite Genomics Group
          The Institute for Genomic Research
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-530-9319
          Fax: 301-838-0208
          Email: carlton@tigr.org
          For clone info, please contact the Malaria Research and Reference
          Reagent Resource Center, ATCC
          http://www.malaria.mr4.org/mr4pages/index.html
          Seq primer: ADP.

FEATURES
     source
     1..696
     /organism="Plasmodium yoelii yoelii"
     /mol_type="mRNA"
     /strain="17XL"
     /sub_species="yoelii"
     /db_xref="taxon:73239"
     /clone="PYCKS34"
     /dev_stage="Asexual blood stages"
     /lab_host="E. coli XL-1 Blue"
     /clone_lib="PyBS"
     /notes="Vector: pAD-GAL4; At 20-25% parasitemia, blood was
collected from BALB/cBYJ mice infected with Py17XL
parasites, and leukocytes removed by passage over
microcrystalline cellulose columns. Total RNA was
isolated using the guanidinium isothiocyanate method, and
mRNA isolated using oligo(dT)-cellulose chromatography.
First strand cDNA synthesis was completed using a 50-base
primer and reverse transcriptase in the presence of
5-methyl dCTP. After second strand synthesis, uneven
termini were treated with Pfu DNA polymerase and EcoRI
adaptors ligated to the blunt ends. The sample was cleaved
with XhoI and separated on a Sephacryl S-500 column.
Size-fractionated cDNA was precipitated and ligated to
HybridZAP arms directionally using EcoRI-XhoI cleaved arms.
After packaging, the phagemid vector (pAD-GAL4) was
excised from the HybridZAP vector and plasmid DNA
isolated."

ORIGIN
Alignment Scores:
Pred. No.:      0.322      Length:      696
Score:          105.50     Matches:      58
Percent Similarity: 41.5%   Conservative: 49
Best Local Similarity: 22.5% Mismatches:    94
Query Match:     6.7%     Indels:      57
DB:              2        Gaps:       12

US-10-734-719-12 (1-291) x BM162427 (1-696)

```

```

Qy      26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGly----- 42
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      9 GAGGCCCTTTTATACAAATTTACTAAATGAGGAAGAGAACTATTTCTGGGATATTAAAT 68
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      43 ---LysLysCysLysAlaValPheTyrThrProGlyPhePhePheGluGlnTyrTyrThr 61
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      69 AAGGAAAAAAGAGAGAAATGAAATGCTGCTGTTGATTTTTTTTGATGATTAATATATCT 128
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      62 LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn--- 80
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      129 GATGATGTTATTTTACAAATTAATAATATGAAATAAAATAGATAATAATTCAAGTTTG 188
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      81 TyrAsnGlnAlaHisLeuGluAsnGlnPheValLys---ThrPheTyrAspTyrPhe 99
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      189 TTAATAATAATAATAATAATAATAAGAGAAATCAATAATAAGACACAGAAAAAACTAAATTA 248
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      100 ProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu---PheAsnAlaTyr 118
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      249 AGCAATTCAGAGATAAAGAGAAATATATCTTTAAATCTACAAAAGAAATGTTATATCATAT 308
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      119 Phe-----LysPheHisGluIleTyrPheAsnGlnArg-----IleThrSer 132
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      309 TATGTTTCAATTCCTAACATTATAAATTATTTTCAACTATAAAATGTTCTTAACAAT 368
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      133 GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGly 152
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      369 GATATATATAAATATTTTATA-----AATAAGAAATTTCTTTATCAGAT 413
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      153 IleAspPheTyrGlnAsnGlySerTyrAlaPheAspThrLys-----Gln 168
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      414 GGTGATAATTATATTTACATGTCAGATCAAAATATATAATTAATGAATGTAATTTTGAA 473
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      169 GluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis 188
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      474 AAAAATATTTTGAAAAAATTTATTTAAAAAGAAATCGAAAAAATATATCTATAAACAT 533
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      189 SerLysAsnThrAspIleLysAlaLeuGluPheLysGluLysThrTyrLysIleLysLeu 208
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      534 GTAAAAAAT-----AATTCAGAAAAATCAAAATA 563
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      209 TyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsn 228
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      564 -----AATAAACTATCAGTCAT-----CAAAATAGAT 590
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      229 SerAsnPheIleIleGlnGluLysAsnTyrThrLysAspIleLeuIleProSerSer 248
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      591 AAAAATTTACCTAATTAAGGAAATCTAATATAAATTTTGGAT----- 632
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Qy      249 GluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleLys 266
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||
Db      633 -----CGAATAATAAGAAAAAATAAAAAAATAAAAAA 662
      |||::: ||| ::::||| ::::||| ::::||| ::::||| ::::||| ::::|||

RESULT 12
LOCUS   DV143930
DEFINITION DV03135A1G11.f1 CV03-normalized library Euphorbia esula cDNA clone
CV03135A1G11.f1 5, mRNA sequence.
ACCESSION DV143930
VERSION   1
KEYWORDS  EST.
SOURCE    Euphorbia esula (leafy spurge)
ORGANISM  Euphorbia esula
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
          rosids; eurosids I; Malpighiales; Euphorbiaceae; Euphorbioideae;
          Euphorbiaeae; Euphorbia.
          1 (bases 1 to 820)
REFERENCE Anderson,J.V., Horvath,D.P., Thimmapuram,J., Liu,L., Hernandez,A.,
          Kim,W.R. and Mikel,M.
AUTHORS   Direct submission, Anderson,J.V. 2005
TITLE     Unpublished (2005)
JOURNAL
COMMENT   Contact: James V. Anderson

```



USDA/ARS  
1605 Albrecht Blvd, Biosciences Research Laboratory, Fargo, ND  
58105, USA  
Tel: 701 239 1263  
Fax: 701 239 1252  
Email: andersjv@fargo.ars.usda.gov  
Adaptors and tags in 5'-end sequenced clones:  
(Vector). TAGCTTGATATCG(End Vector)(Start  
EcoRI adaptor)AATCCATGTTGTTGG (End EcoRI adaptor) (Start  
Insert). . . AAAAAAAAAAAAAAAAAA(End Insert) (Start Tag/TCGGT(End  
Tag)(Start NotI site/Vector)GGCGCCGACCCGCG. . . Base  
Calling/Quality Scores: PHRED from Washington University Genome  
Center. Vector Trimming: Cross match from Washington University  
Genome Center PHRAP suite. Low quality bases (Phred score < 20)  
were trimmed from both ends of the sequence by an in-house script.  
This sequence is vector free and at least 200 bp in length.

## PCR Primers

FORWARD: TAATACGACTCACTATAGG(T7)

BACKWARD: ATTAACCTCACTAAAG(T3)

Insert Length: 820 Std Error: 0.00

Plate: CV03135A1 row: G column: 11

Seq primer: TAATACGACTCACTATAGG (T7)

High quality sequence stop: 820.

## FEATURES

source

1. 820 /organism="Euphorbia esula"

/mol\_type="mRNA"

/db\_xref="taxon:3993"

/clone="CV03135A1G11.f1"

/lab\_host="DH10B"

/clone\_lib="CV03-normalized library"

/note="Vector: pBS II SK(+); Site 1: EcoRI(5' side of

insert); Site 2: NotI (3' side of insert); Tissue:

Seasonal crown buds from outdoor plants, whole plant  
tissue (leaf, stem, flowers, seeds, shoots, galls, root  
and crown buds) from outdoor plants, cold acclimated  
tissues (leaf, stem, and crown buds), crown and root buds  
from decapitated greenhouse p ants, and seedlings. The  
library was constructed and normalized as described by  
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
Research 6(9): 791-806. An identifying tag was added at the  
3'during cDNA synthesis: insertAAAAAAAAAAAAAAAAATGCGT."

## ORIGIN

## Alignment Scores:

Pred. No.: 0.455 Length: 820  
Score: 105.00 Matches: 54  
Percent Similarity: 38.6% Conservative: 46  
Best Local Similarity: 20.8% Mismatches: 99  
Query Match: 6.7% Indels: 60  
DB: 10 Gaps: 11

US-10-734-719-12 (1-291) x DV143930 (1-820)

Qy 48 ValPheTyrThrProGlyPhePheGluGlnTyrTyrThrLeuLysHisLeuGln 67  
Db 49 ATCTTCGATTCGATGGTATGCTTATGAAACTTCGTTGCATTC----- 93  
Qy 68 AsnGlnGluTyrGluThrGluLeuLeuMet-----CysSerAsnTyrAsnGlnAlaHis 85  
Db 94 -----GTTGCAATACCTTAATTCCTATGACTGCTCCAAATCAAAATCCAC 144  
Qy 86 LeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGly 105  
Db 145 CTAAAA-----TATAAC-----ATTCGCGATCGGAAGGATCCC 177  
Qy 106 TyrAspPhePheLysGlnLeu-----LysGluPheAsnAlaTyrPheLysPheHis 122  
Db 178 GAGTATTTTTCGAAATTCCTTGATTGCAACAAACAAACAAATCAAAATACACGTTTCGG 237  
Qy 123 GluLeuTyrPheAsnGln-----ArgIleThrSer----- 132  
Db 238 GAGTTTCGTTTTCACAAATGCTTCATTTGTTAATTAACAAACCATTAATTCGCAATTTATG 297

Qy 133 -----GlyVal-----TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148  
Db 298 CTCGTGGAAATGTAAATTTGGAGATGTCACATCACTGATATATTCATATGCAATCG 357  
Qy 149 TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGln 168  
Db 358 AATGATCAAGCAATAGATATATTCCTATCTGGAAGTCGCTGCTCAAAATCCTTAAAAATTA 417  
Qy 169 GluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis 188  
Db 418 AGCAATTTGTCGAAGCTT-----AAAAGCGTGGTTTACT 453  
Qy 189 SerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeu 208  
Db 454 TCCAAATCTTGGAAATATGAGTATGAGATATCTTCGCCGGATTTCTATCTTGAAT 513  
Qy 209 TyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsn 228  
Db 514 TCATGC-----CGAAGCTTAAAG 531  
Qy 229 SerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSer 248  
Db 532 AAATTCGAGCTTATCATAGTTGTTTCCAAATGTCAAAACAAATTAATTCGCTGAATTTGCCATCT 591  
Qy 249 GluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIle-----LysIleLys 266  
Db 592 TCACTTTGGTACTCTTATGATACCATCTTAATTTGGAAGATGATGATGAAAGTTCC 651  
Qy 267 GluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeuLeuProSerAspIleLys 285  
Db 652 AAAGAAATTCATAGAGTTTGGTTAAAGACATTCGCAACAGCTTCATCATATCAAG 708  
RESULT 13  
CL024751/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source  
1. .1130 /organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-20C18"  
/sex="male"  
/cell\_line="Stock 248 F7A2, inbred N7"  
/clone\_lib="CH216"  
/notes="Vector: pTARBAC3.1; CHORI-216 Xenopus tropicalis  
BAC library"

```
Alignment Scores:
Pred. No.: 0.791 Length: 1130
Score: 104.50 Matches: 71
Percent Similarity: 34.2% Conservative: 38
Best Local Similarity: 22.3% Mismatches: 93
Query Match: 6.7% Indels: 117
DB: 15

US-10-734-719-12 (1-291) x CL024751 (1-1130)
Qy 2 LysLysValIleIleAlaGly-----AsnGlyProSerLeuLysGlu 15
Db 1051 AGAAGAGTTTATTATTATAGGAGAGAGTGTCTTTTCTTTTCTGAGGTCCA----- 1004
Qy 16 IleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPhe 35
Db 1003 -----AAAATTCACCCAGCTTTTGTATATACATGGGCTTTTCAACAAAGGTTT--- 959
Qy 36 GluAspLysTyrTyrLeuGlyLysCysLysAlaValPheTyrThrProGlyPhePhe 55
Db 958 -----AAAAGATTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTTT 905
Qy 56 -----PheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThr 73
Db 904 CCAATTTTCAACCAAGACACTGCACAAAAAGGTAATTTGGGCTTTTGAAGTGGAAAAA 845
Qy 74 GluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLys 93
Db 844 GTACAATTGCCCTAGGAAACTGGA--- 818
Qy 94 ThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLys 113
Db 817 -----GGTCTCGTAGTCTTTTAAAGAAATACCA 791
Qy 114 GluPheAsnAlaTyrPheLysPhe-HisGluIleTyrPheAsnGlnArgIleThrSerGl 133
Db 790 AACATCAGGGATATTTAGATTATACATATAGTATGCTGCATTAACTGTACAGCGA 731
Qy 133 yValTyrMetCysAlaValAlaIle----- 141
Db 730 TTTTGTGTTTCTGCTGGTGAATATTGTAATAAGTATTGCCTTAGTTGGGGTTACTT 671
Qy 142 -----AlaLeuGlyTyrLysGluIleTyr-----LeuSerGlyIleAspPheTyrGl 157
Db 670 CTGGACAGGAACCTGCTGGGTACACACATATTTGGTATCGTTGGAATT----- 622
Qy 157 nAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAs 177
Db 621 -----GGAGTATCAGGCTTTTACAAACAATAAAAAAAGTGAATTAATTAATTTCT 566
Qy 177 pPheLysAsnAspArgSer-----HisTyrIleGl 187
Db 565 ATGGAATAAAACCTCAAAATATACAGAAATTTTTCATAATTTTATTTTTCATATT 506
Qy 187 yHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLy 207
Db 505 TCACCAAAATATACACA-----TCATATCTCCAGAAAGTTATATAAATTTG 461
Qy 207 sLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLe 227
Db 460 GTATGTATGTCTGAAGCCCA----- 442
Qy 227 uAsnSerAsnPheIleIleGlnGluLysAsnAsn-TyrThr-----LysAspI 243
Db 441 -----ATTAGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 395
Qy 243 leLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysI 263
Db 394 TTTTCTTACCA-----AAAAACAATT-----G 374
Qy 263 leLysIleLysGluAsnValTyrTyrLysIleIleLysAspLeuLeuArgLeu 280
Db 280 ----- 280
```

```
Db 373 TTAAAGTGAATGAG-----TACAAATGCTTAAAAAACGTTGCACTG 330

RESULT 14
AY411754 1131 bp DNA linear GSS 16-DEC-2003
LOCUS Pan troglodytes SIAT9E gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY411754
VERSION AY411754.1 GI:39767722
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
REFERENCE 1 (bases 1 to 1131)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1131)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1131
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1131
/genes="SIAT9E"
/locus_tag="HCM4329"
gene
ORIGIN
Alignment Scores:
Pred. No.: 0.792 Length: 1131
Score: 104.50 Matches: 51
Percent Similarity: 39.5% Conservative: 30
Best Local Similarity: 24.9% Mismatches: 89
Query Match: 6.7% Indels: 35
DB: 14

US-10-734-719-12 (1-291) x AY411754 (1-1131)
Qy 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
Db 484 AAGAAAGTGTCTGTGTAGTGGGCAACGAGGAGCATTTGAAGAAACAGCCGCTGTGGAGGAG 543
Qy 22 ProAsnAspPheAsp-----ValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 544 ATCAACAGCCCGACTTCGTCTTCCGGTGCACCTGCCCCCATCTCAGAGAAGTACACC 603
Qy 41 Leu-----GlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGln 58
Db 604 ATGGACGTGGGGGTGAAGACGAGTGTGTCTACTGTGAACCCCGCATCATCACAGAGAGG 663
Qy 59 TyrTyrThrLeuLys-----HisLeuIleGlnAsnGlnGluTyr 71
Db 664 TTCACNNGTGGAGAGAGTGGCGGCGCGCTTCTATCTCGGTGTCGACGTGTACGAG--- 720
Qy 72 GluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPhe 91
Db 721 AACGCGTGTGTGTGTCTGCTGCTCTTCTACACACGCGCAACCGAGCTGTCCATCCGC 780
```

QY 92 VallysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGln 111  
Db 781 GTCAAAATCGTCTCGAGACACTCGAATCGCGCAAGCTGTCTACTACTTCATCGCGAG 840  
QY 112 ---LeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgile 130  
Db 841 TACCTGGTCAACGTCTGGCTACTGGCTACGCTGGGGGTGGC---CCCAAGCGCATC 897  
QY 131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150  
Db 898 AGCACCGGCTCATCTGTGCTACTCGCGGCTGGAGCTC---TGTGAGGAGGTGCACCTC 954  
QY 151 SerGlyIleAspPheTyrGlnAsnGlySerTyrAlaPheAspThrLysGlnGluAsn 170  
Db 955 TTGGCTTC-----TGGGCTTCCCATGAACCCCTCGGCG 990  
QY 171 LeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLys 190  
Db 991 CTC-----TACATCACTCACCACTAC 1011  
QY 191 AsnThrAspIleLys 195  
Db 1012 TATGACAAGTCAAG 1026

RESULT 15  
CX373647  
LOCUS  
DEFINITION JGI XZT57939.fwd NIH XGC tropTad5 Xenopus tropicalis cDNA clone  
IMAGE:7632400 5', mRNA sequence.

ACCESSION CX373647.2 GI:74296874  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Xenopus tropicalis (western clawed frog)  
Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
1 (bases 1 to 896)  
Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,  
Brokstein,P. and Lindquist,E.A.  
DOE Joint Genome Institute Xenopus tropicalis EST project  
Unpublished (2004)  
On Jan 5, 2005 this sequence version replaced gi:57142204.  
Other ESTs: JGI XZT57939.rev  
Contact: Lindquist,E.A., Richardson,P.  
DOE Joint Genome Institute  
2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Tel: 925 296 5600  
Fax: 925 296 5710  
Email: cdna@jgi-psf.org

Tissue Procurement: Richard M. Harland Laboratory, University of  
California, Berkeley: <http://tropicalis.berkeley.edu/home>  
cDNA Library Preparation: Richard M. Harland Laboratory, University  
of California, Berkeley  
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>  
Clone Distribution: I.M.A.G.E. Consortium/LLNL:  
<http://image.llnl.gov>  
Naming Conventions: EST name is generated by the concatenation of  
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'  
indicates a forward sequencing read of the insert. It does not  
necessarily reflect the orientation of the insert.  
Plate: XZT 0601 row: f column: 14  
High quality sequence stop: 806.

FEATURES  
source  
1..896  
/organism="Xenopus tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="IMAGE:7632400"  
/tissue\_type="whole embryo"  
/dev\_stage="Tadpole (St. 36-41)"  
/lab\_host="E. coli XLI-Blue derivative, Stratagene

ElectroTen-Blue"  
/clone lib="NIH XGC tropTad5"  
/note="Vector: pCS108; Site 1: Sali; Site 2: NotI; Tadpole  
library constructed by Russell B. Fletcher in R. Harland's  
lab using poly A RNA and oligo dt primers (invitrogen  
SuperScript Plasmid System for cDNA Synthesis and  
Cloning). Sali (5' end) -NotI (3' end) cDNA was inserted  
into vector pCS108  
(<http://mcb.berkeley.edu/labs/harland/pages/plasmids.html>)  
."

ORIGIN

Alignment Scores:  
Pred. No.: 0.664 Length: 896  
Score: 104.00 Matches: 54  
Percent Similarity: 36.1% Conservative: 25  
Best Local Similarity: 24.7% Mismatches: 76  
Query Match: 6.7% Indels: 64  
DB: 9 Gaps: 11

US-10-734-719-12 (1-291) x CX373647 (1-896)

QY 10 GlyProSerLeuLysGluIleAspTyrSerArgLeuProAsn---AspPheAspValPhe 28  
Db 173 GGCTACTTAATAAAATTCGACAAATTCACAACTGCCTGAGGTAGAACTAGACTTTT 232  
QY 29 ArgCysAsnGlnPheTyrPheGlu-----AspLysTyrTyrLeuGlyLysLysCys 45  
Db 233 GAAATCCTGCGCTGTTTTTCCCGAGTATACGATCAAAATACCTGATGAAAGCTGC 292  
QY 46 LysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyrThrLeuLysHisLeu 65  
Db 293 AAA-----AATCTCAAGGTGGA 310  
QY 66 IleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHis 85  
Db 311 TTGCAGGAAGTGCAGAGCAGCTGGAACCTCAAGCGTATTGGTCTCTCAGCATCAAGCA--- 367  
QY 86 LeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGly 105  
Db 368 -----GGC 370  
QY 106 TyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyr 125  
Db 371 TCTGAC-----TCATTACTTACTGGAATGGCATCTTTAAATGAGAGAGATGTTTC 421  
QY 126 PheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyr 145  
Db 422 TTGAGAGATCACATAGATGATGCAAAATAC---TGTGGCCATTATATGGTCTGGGA--- 475  
QY 146 LysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAsp 165  
Db 476 -----TCCGGATCGTCTTATGTACAGAAATGTACAGGAATGCAATATGAA 520  
QY 166 ThrLys-----GlnGluAsnLeuLysLeuAlaProAspPhe 178  
Db 521 GAGAGACCAATAAACAATCATGACACCAAGAGAAACTGCACCTTTTATGCCAAACTAC 580  
QY 179 LysAsnAspArgSerHisTyrIleGly---HisSerLysAsnThrAspIleLysAlaLeu 197  
Db 581 AGAAATTTGATGATGTTTATCTCTGGTGATATCTCTGGACACTGAATTTCCCCCATCT 640  
QY 198 GluPheLeuGluLysThrTyrLysIleLysLeu-----CysProAsn 214  
Db 641 CCCTGCACACAG-----TTGTCTGCTTCATCTTGTATCAAT 679

Search completed: August 21, 2006, 20:38:49  
Job time : 6310 secs

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 15:00:51 ; Search time 176.5 Seconds  
(without alignments)  
4627.416 Million cell updates/sec

Title: US-10-734-719-12  
Perfect score: 1563  
Sequence: 1 MKKVIIAGNPSLKEIDYSR.....KLIKDLLRLPSDKHYFKGK 291

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/abes/ABSSWEB spool/US10734719/runat 21082006 141818 12415/app query.fasta\_1  
-DB=Issued Patents NA -QFMT=fastap -SURFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs05p -USER=US10734719 @CGN 1.1 381 @runat 21082006 141818 12415  
-NCPU=6 -ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS COMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP COMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	873	US-09-816-028A-11	Sequence 11, Appl
2	1563	100.0	873	US-10-303-162-11	Sequence 11, Appl
3	1563	100.0	873	US-10-303-134-11	Sequence 11, Appl
4	1563	100.0	873	US-10-303-118-11	Sequence 11, Appl
5	1563	100.0	873	US-10-303-128-11	Sequence 11, Appl
6	1563	100.0	873	US-10-735-419-11	Sequence 11, Appl
7	1557	99.6	876	US-09-495-406-8	Sequence 8, Appl
8	1557	99.6	876	US-09-816-028A-8	Sequence 8, Appl

9	1557	99.6	876	3	US-10-303-162-8	Sequence 8, Appl
10	1557	99.6	876	3	US-10-303-134-8	Sequence 8, Appl
11	1557	99.6	876	3	US-10-303-118-8	Sequence 8, Appl
12	1557	99.6	876	3	US-10-303-128-8	Sequence 8, Appl
13	1557	99.6	876	5	US-10-735-419-8	Sequence 8, Appl
14	1553	99.4	873	3	US-09-816-028A-13	Sequence 13, Appl
15	1553	99.4	873	3	US-10-303-162-13	Sequence 13, Appl
16	1553	99.4	873	3	US-10-303-134-13	Sequence 13, Appl
17	1553	99.4	873	3	US-10-303-118-13	Sequence 13, Appl
18	1553	99.4	873	3	US-10-303-128-13	Sequence 13, Appl
19	1553	99.4	873	5	US-10-735-419-13	Sequence 13, Appl
20	1534	98.1	876	3	US-09-495-406-4	Sequence 4, Appl
21	1534	98.1	876	3	US-09-816-028A-4	Sequence 4, Appl
22	1534	98.1	876	3	US-10-303-162-4	Sequence 4, Appl
23	1534	98.1	876	3	US-10-303-134-4	Sequence 4, Appl
24	1534	98.1	876	3	US-10-303-118-4	Sequence 4, Appl
25	1534	98.1	876	3	US-10-303-128-4	Sequence 4, Appl
26	1534	98.1	876	5	US-10-735-419-4	Sequence 4, Appl
27	1524	97.5	876	3	US-09-495-406-2	Sequence 2, Appl
28	1524	97.5	876	3	US-09-816-028A-2	Sequence 2, Appl
29	1524	97.5	876	3	US-10-303-162-2	Sequence 2, Appl
30	1524	97.5	876	3	US-10-303-134-2	Sequence 2, Appl
31	1524	97.5	876	3	US-10-303-118-2	Sequence 2, Appl
32	1524	97.5	876	3	US-10-303-128-2	Sequence 2, Appl
33	1524	97.5	876	5	US-10-735-419-2	Sequence 1, Appl
34	1524	97.5	11474	3	US-09-495-406-1	Sequence 1, Appl
35	1524	97.5	11474	3	US-09-816-028A-1	Sequence 1, Appl
36	1524	97.5	11474	3	US-10-303-162-1	Sequence 1, Appl
37	1524	97.5	11474	3	US-10-303-134-1	Sequence 1, Appl
38	1524	97.5	11474	3	US-10-303-118-1	Sequence 1, Appl
39	1524	97.5	11474	3	US-10-303-128-1	Sequence 1, Appl
40	1524	97.5	11474	5	US-10-735-419-1	Sequence 1, Appl
41	1513	96.8	876	3	US-09-495-406-6	Sequence 6, Appl
42	1513	96.8	876	3	US-09-816-028A-6	Sequence 6, Appl
43	1513	96.8	876	3	US-10-303-162-6	Sequence 6, Appl
44	1513	96.8	876	3	US-10-303-134-6	Sequence 6, Appl
45	1513	96.8	876	3	US-10-303-118-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-09-816-028A-11  
; Sequence 11, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-09-816-028A-11

Alignment Scores: 3.62e-174 Length: 873  
Pred. No.: 3.62e-174

```
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x US-09-816-028A-11 (1-873)

QY 1 MetLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGCGAAATGGCAAGTTTAAAGAAATGATTATTTCAAGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGGATGATTAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 120

QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAAAGCAGTGTTTACACCCCTGGTTCTCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGCAACTAATATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCAGCAAAATTTATTTCAATCAAGAAATTAATGAGGGGTCTATATGTCAGTAGGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyLysAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTCGATTTTAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAAATAAAACTATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTT 660

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAAGTACAGCCCAATTTAAATTCAAATTTTATCATACAGAAAAATACTACTACT 720

QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780

QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAGATTA 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAAGCATTTATTTCAAAAGAAAA 873
```

## RESULT 2

```
US-10-303-162-11
; Sequence 11, Application US/10303162
; Patent No.: 6723545
; GENERAL INFORMATION:
```

```
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-162-11
```

## Alignment Scores:

```
Pred. No.: 3,62e-174 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-12 (1-291) x US-10-303-162-11 (1-873)

```
QY 1 MetLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGCGAAATGGCAAGTTTAAAGAAATGATTATTTCAAGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGGATGATTAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 120

QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAAAGCAGTGTTTACACCCCTGGTTCTCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGCAACTAATATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCAGCAAAATTTATTTCAATCAAGAAATTAATGAGGGGTCTATATGTCAGTAGGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyLysAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTCGATTTTAAAT 540
```

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerIleLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTCTTAAACAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnLeuTyrThr 240  
DB 661 ATAGAACTAGCGCAAAATTTAAATTTTATCATCAAGAAATAAATACTACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATATTAATTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAAGCAATTATTTCAAAAGGAAAA 873

## RESULT 3

US-10-303-134-11  
; Sequence 11, Application US/10303134  
; Patent No. 6825019  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-134-11

## Alignment Scores:

Pred. No.: 3,62e-174 Length: 873  
Score: 1563.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-134-11 (1-873)

QY 1 MeLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTTATTATTGCTGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAAAATGATTTTGTATGTTATTAGATGTAATCAATTTTATTTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePhePheGluGlnTyrTyr 60  
DB 121 CTTGGTAAAAATGCAAGCAGTGTCTTTTACACCCCTGGTCTTCTTTGAGCAACTACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATAGAGCCGAACTAAATATATGTGTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATATTATTTCTCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCATTTGGGATATGATTTTNTAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyIleAspPheTyrGlnAsnGlySer 140  
DB 361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTTATGCTTTTGTATCAACAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTCTTAAACAGTCTTTTAGCAAAATTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
DB 661 ATAGAACTAGCGCAAAATTTAAATTTTCAAAATTTTATCATCAAGAAATAAATACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATATTAATTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAAGCAATTATTTCAAAAGGAAAA 873

## RESULT 4

US-10-303-118-11  
; Sequence 11, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-118-11

Alignment Scores:
Pred. No.: 3,62e-174 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-118-11 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGGATGTTATAGATGTAATCAATTTTATTGTAAGATAAAATACAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCCTTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACGCACTAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAGAACTTTTACGATTTATTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCCAGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCATTATGCTTTTGATACCAACAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATAAAACTATATTGCTTATGTCCTCAACAGCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAAATACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTTCTAGTAGGCGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAGTTTATATAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 5
US-10-303-128-11
; Sequence 11, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-128-11

Alignment Scores:
Pred. No.: 3,62e-174 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-128-11 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGGATGTTATAGATGTAATCAATTTTATTGTAAGATAAAATACAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCCTTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACGCACTAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAGAACTTTTACGATTTATTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCCAGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCATTATGCTTTTGATACCAACAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATAAAACTATATTGCTTATGTCCTCAACAGCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAAATACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTTCTAGTAGGCGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780
```

```
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnAiqIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATATGTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTAAATTTTATCATACAAGAAAAATTAACACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATCTCTTAGTAGGCTTATGGAATAATTTTCAAAAAATATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuAiqLeu 280
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTACAGTTGATAAAAGATCTATTAAAGATT 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 6
US-10-735-419-11
; Sequence 11, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-735-419-11
```

Alignment Scores:  
Pred. No.: 3,62e-174 Length: 873  
Score: 1563.00 Matches: 291

```
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-734-719-12 (1-291) x US-10-735-419-11 (1-873)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGTATGATGATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACCTAATATATGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAATGCTTATTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACAGAAATTTATTTCAATCAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATTAACACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATCTCTTAGTAGGCTTATGGAATAATTTTCAAAAAATATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuAiqLeu 280
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTACAGTTGATAAAAGATCTATTAAAGATT 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 7
US-09-495-406-8
; Sequence 8, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
```





```
Db 181 ACTTTAAACATTTAAATCCAAAATCAAGAAATATGAGACCGAAGCTAAATATGTGTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValIysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTAATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAGAAATTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGCTATATGTGTGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAAATAAACTATATTTGCTTATGTCTTAATAGTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTTAAATTTCAAAATTTTATCATACAAGAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTCTAGTAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
```

## RESULT 9

```
US-10-303-162-8
; Sequence 8, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-162-8
```

## Alignment Scores:

```
Pred. No.: 1.85e-173 Length: 876
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-12 (1-291) x US-10-303-162-8 (1-876)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTTATTTGCTGGAAATGACCAAGCTTTAAAAAGAAATTTGATTATTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAAGCTAAATATGTGTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTAATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAGAAATTTTAAATTTTAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTATATATGTGTGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAAATAAACTATATTTGCTTATGTCTTAATAGTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTTAAATTTCAAAATTTTATCATACAAGAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
```

```
Db      841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
RESULT 10
US-10-303-134-8
; Sequence 8, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-134-8
Alignment Scores:
Pred. No.: 1,85e-173 Length: 876
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 3 Gaps: 0
US-10-734-719-12 (1-291) x US-10-303-134-8 (1-876)
QY      1 MetLysLysValIlellelaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1 ATCAAAAAAGTTATTATTCGCGAAATGACCAAGTTTAAAGAAAAATTCATTATTCAAG 60
QY      21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61 CTACCAAAATGATTTTGTATGATTATAGATGTAATCAATTTTATTGAAAGATAAATACTAT 120
QY      41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      121 CTGTGTAATAAATGCAACAGCTGTTTACACCCCTAATTCCTCTTGAGCAATACTAC 180
QY      61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGCAACTAATATTATGTTCTAAT 240
QY      81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY      101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAGAAATTTAATGCTTATTTTAA 360
QY      121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361 TTTACAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGTCAGTAGCC 420
QY      141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCGATTTTATCAAAATCGGTCA 480
QY      161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db      481 TCTTATGCTTTTGATACCAAAACAAGAAAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540
QY      181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY      201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601 GAAAAAACTTTACAAAATAAAACTATATTGCTTATGTCTCTAATAGTCTTTTAGCAAAATTT 660
QY      221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db      661 ATAGAACTAGCGCAAAATTTAAATTTTCAATTTTATCATACAAAGAAAAATAACTACACT 720
QY      241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721 AAAGATATATCATACCTTCTAGTAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
QY      261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY      281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
RESULT 11
US-10-303-118-8
; Sequence 8, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-118-8
Alignment Scores:
Pred. No.: 1,85e-173 Length: 876
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 3 Gaps: 0
US-10-734-719-12 (1-291) x US-10-303-118-8 (1-876)
QY      1 MetLysLysValIlellelaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1 ATCAAAAAAGTTATTATTCGCGAAATGACCAAGTTTAAAGAAAAATTCATTATTCAAG 60
QY      21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61 CTACCAAAATGATTTTGTATGATTATAGATGTAATCAATTTTATTGAAAGATAAATACTAT 120
QY      41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db      121 CTGTGTAATAAATGCAACAGCTGTTTACACCCCTAATTCCTCTTGAGCAATACTAC 180
QY      61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGCAACTAATATTATGTTCTAAT 240
QY      81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY      101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAGAAATTTAATGCTTATTTTAA 360
QY      121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361 TTTACAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGTCAGTAGCC 420
QY      141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
```

```
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTGATGTGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATAATATGTTCTTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGTCTATATGTTGCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyLeAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACATATTCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
QY 201 GlyLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAACTTACAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCTA 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTCATACCTCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTACAGTTGATTAAGATCTATTAAGATT 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873
```

## RESULT 12

```
US-10-303-128-8
; Sequence 8, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
```

```
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-128-8
```

## Alignment Scores:

Pred. No.:	1,85e-173	Length:	876
Score:	1557.00	Matches:	290
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	3	Gaps:	0

US-10-734-719-12 (1-291) x US-10-303-128-8 (1-876)

```
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTGATGTATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATAATATGTTCTTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGTCTATATGTTGCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACATATTCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
QY 201 GlyLysThrTyrLysLysLeuLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAACTTACAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCTA 660
```

```
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTAGGCTTATGGAAATTTTCAAAAAATATTAAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTATTTTCAAGGAAAA 873

RESULT 13
US-10-735-419-8
; Sequence 8, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-735-419-8

Alignment Scores:
Pred. No.: 1,85e-173 Length: 876
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 5 Gaps: 0

US-10-734-719-12 (1-291) x US-10-735-419-8 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATCTGCTGGAATGACCAGTTTAAAGAAATTTGATTTATTTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTGATGTAATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGACAGTGTTTATACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAAAACATTTAATCCAAAAATCAAGAATATGAGACCGAAGCACTAATATGTGTTCTAAT 240
```

```
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAITTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTTCAATCAAGAAATTAACCTCAGGGGCTTATATGTGTGCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGTATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAAATTTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTCAATTTTATCATACAAAGAAAAATACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTATTTTCAAGGAAAA 873

RESULT 14
US-09-816-028A-13
; Sequence 13, Application US/09816028A
; Patent No. 6695705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
```

## US-09-816-028A-13

Alignment Scores:  
Pred. No.: 5,43e-173 Length: 873  
Score: 1553.00 Matches: 289  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 2  
Query Match: 99.4% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x US-09-816-028A-13 (1-873)

```
Qy 1 MetLysLysValIleLeuAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTATTGGAAGATAAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTATCCAAATCAAGATATGAGACCAACTAATATTATGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTTATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAAAACAATTAAGAAATTTTAAATGCTTTTAAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCATCAAGAAATTAATCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATACAAATATAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATATAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873
```

RESULT 15

## US-10-303-162-13

; Sequence 13, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-162-13

## Alignment Scores:

Pred. No.: 5,43e-173 Length: 873  
Score: 1553.00 Matches: 289  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 2  
Query Match: 99.4% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-162-13 (1-873)

```
Qy 1 MetLysLysValIleLeuAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTATTGAAAGATAAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTATCCAAATCAAGATATGAGACCAACTAATATTATGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTTATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAAAACAATTAAGAAATTTTAAATGCTTTTAAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCATCAAGAAATTAATCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATACAAATATAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATATAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873
```

```
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db |||||
481 TCTTAAGCTTTGATACCAACAGAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db |||||
541 GATCGCTCACACTATATCGGACATAGTAAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db |||||
601 GAAAAAACTTACAAAATAAAACTATATTGCTTATGTCTTAATAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db |||||
661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATCAAGAAAAAATAACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db |||||
721 AAAGATATACTCATACCTTCTTAGTGAGGCTTATGGAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db |||||
781 AAAAAAATAAAAAATTAAGAAAATGTTTATTACAGTTTGATAAAGATCTATTAAAGATTAA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db |||||
841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
```

Search completed: August 21, 2006, 17:44:40  
Job time : 182.5 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 15:32:54 ; Search time 1385 Seconds  
(without alignments)  
3872.601 Million cell updates/sec

Title: US-10-734-719-12  
Perfect score: 1563  
Sequence: 1 MKKVIIAGNPSLKEIDYSR.....KLIKDLRLPSDIKHVFKGK 291

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DBV=xlh  
-Q=/abs/ABSSWEB.spool/US10734719/runat.21082006.141829.12577/app\_query.fasta\_1  
-DB=Published Applications NA Main -OFMT=fastap -SUFFIX=P2N.rnpbm  
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h  
-USER=US10734719 @CGN 1.1 2326 @runat.21082006.141829.12577 -NCPU=6 -ICPU=3  
-NO\_MMAP -NEG\_SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database : Published Applications NA Main.\*

```

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	873	3	US-09-816-028A-11

## ALIGNMENTS

## RESULT 1

```

US-09-816-028A-11
; Sequence 11, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS

```

2	1563	100.0	873	7	US-10-303-161-11	Sequence 11, Appl
3	1563	100.0	873	7	US-10-303-118-11	Sequence 11, Appl
4	1563	100.0	873	7	US-10-303-128-11	Sequence 11, Appl
5	1563	100.0	873	7	US-10-303-134-11	Sequence 11, Appl
6	1563	100.0	873	7	US-10-303-162-11	Sequence 11, Appl
7	1563	100.0	873	9	US-10-735-419-11	Sequence 11, Appl
8	1563	100.0	873	9	US-10-820-536-11	Sequence 11, Appl
9	1563	100.0	873	9	US-10-845-408-11	Sequence 11, Appl
10	1563	100.0	873	9	US-10-845-412-11	Sequence 11, Appl
11	1563	100.0	873	9	US-10-846-219-11	Sequence 11, Appl
12	1563	100.0	873	9	US-10-821-604-11	Sequence 11, Appl
13	1563	100.0	873	9	US-10-847-983-11	Sequence 11, Appl
14	1563	100.0	873	9	US-10-821-573-11	Sequence 11, Appl
15	1563	100.0	873	9	US-10-850-807-11	Sequence 11, Appl
16	1563	100.0	873	9	US-10-850-125-11	Sequence 11, Appl
17	1563	100.0	873	9	US-10-830-825-11	Sequence 11, Appl
18	1563	100.0	873	10	US-10-962-334-11	Sequence 11, Appl
19	1563	100.0	873	10	US-10-830-997-11	Sequence 11, Appl
20	1563	100.0	873	10	US-10-962-235-11	Sequence 11, Appl
21	1563	100.0	873	10	US-10-961-882-11	Sequence 11, Appl
22	1557	99.6	876	3	US-09-816-028A-8	Sequence 8, Appl
23	1557	99.6	876	7	US-10-303-161-8	Sequence 8, Appl
24	1557	99.6	876	7	US-10-303-118-8	Sequence 8, Appl
25	1557	99.6	876	7	US-10-303-128-8	Sequence 8, Appl
26	1557	99.6	876	7	US-10-303-134-8	Sequence 8, Appl
27	1557	99.6	876	7	US-10-303-162-8	Sequence 8, Appl
28	1557	99.6	876	9	US-10-735-419-8	Sequence 8, Appl
29	1557	99.6	876	9	US-10-820-536-8	Sequence 8, Appl
30	1557	99.6	876	9	US-10-845-408-8	Sequence 8, Appl
31	1557	99.6	876	9	US-10-845-412-8	Sequence 8, Appl
32	1557	99.6	876	9	US-10-846-219-8	Sequence 8, Appl
33	1557	99.6	876	9	US-10-821-604-8	Sequence 8, Appl
34	1557	99.6	876	9	US-10-847-983-8	Sequence 8, Appl
35	1557	99.6	876	9	US-10-821-573-8	Sequence 8, Appl
36	1557	99.6	876	9	US-10-850-807-8	Sequence 8, Appl
37	1557	99.6	876	9	US-10-850-125-8	Sequence 8, Appl
38	1557	99.6	876	9	US-10-830-825-8	Sequence 8, Appl
39	1557	99.6	876	10	US-10-962-334-8	Sequence 8, Appl
40	1557	99.6	876	10	US-10-830-997-8	Sequence 8, Appl
41	1557	99.6	876	10	US-10-962-235-8	Sequence 8, Appl
42	1557	99.6	876	10	US-10-961-882-8	Sequence 8, Appl
43	1553	99.4	873	3	US-09-816-028A-13	Sequence 13, Appl
44	1553	99.4	873	7	US-10-303-161-13	Sequence 13, Appl
45	1553	99.4	873	7	US-10-303-118-13	Sequence 13, Appl



```

; LOCATION: (1) .. (873)
; ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; ; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-09-816-028A-11

```

Alignment Scores:		
Pred. No.:	2.58e-152	Length:
Score:	153.00	Matches:
Percent Similarity:	100.0%	Conservative:
Best Local Similarity:	100.0%	Mismatches:
Query Match:	100.0%	Indels:
DB:	3	Gaps:

US-10-734-719-12 (1-291) x US-09-816-028A-11 (1-873)

Qy	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAATAAGCTATATATGTCTGGAAATGGACCAAGTTTAAAGAAATGATTTATTTCAAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAATGATTTTGATGATATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr	60
Db	121	CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTCTTCTTTGAGCAATACTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAACATTTTAATCCAAAAATCAAGAAATATGAGACCGCAACTAAATATGTGTCTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAATAACCTTTTACGATTTATTTCTCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTTAAA	360
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTTCACGAAATTTATTTTCAATCAAAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGATACAAAGAAATTTAICTTTTCGGGAATTGATTTTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCATTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCGCTCACACTATATCGGCACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTTA	600
Qy	201	GluLysThrTyrLysIleLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAACTTTACAAAATAAATACTATTTGCTTATGTCTCTAACAGTCTTTTACCAAAATTTT	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATCAAGAAAAATAAATCACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAATAAAAAATTAAGAAAAATGTTTTATTACAAGTTGATAAAGATCTATTAAAGATTTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTATATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAAATATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAAGAAAAATAACTACACT 720  
Qy 241 LysAspIleLeuLysLeuSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATATCTCATACCTTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATATAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

## RESULT 3

US-10-303-118-11  
; Sequence 11, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-118-11

## Alignment Scores:

Pred. No.:	2 58e-152	Length:	873
Score:	1563.00	Matches:	291
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-734-719-12 (1-291) x US-10-303-118-11 (1-873)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTTATGCTGAAATGGAACCAAGTTTAAAGAAATTTGATTTATCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTGTATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePhePheGluGlnTyrTyr 60  
Db 121 CTGGTAAAAAATGCAAGACAGTGTTTTACACCCCTGGTTCTCTTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
Db 181 ACITTAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCTC 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACAGAAATTTAATTTCAATCAAGAAATTAACCTCAGGGGCTCTATATGTGTCAGTAG 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrIleuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAAATATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATAACTACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATATCTCATACCTTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATATAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

## RESULT 4

US-10-303-128-11  
; Sequence 11, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21

```
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-128-11

Alignment Scores:
Pred. No.: 2,58e-152 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-128-11 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGATGTAATCAATCAATTTTATTGGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyr 60
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTCGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTTAATCCAAATCAAGAAATGAGACCGAACTAATTAIGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAGAACTTTTACGATTATTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTCCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAAAACAGAAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACATATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACCCTTAAAAATAAACTATATTGCTTATGTCTTACACAGTCTTTTAGCAAAATTTT 660

; PRIOR APPLICATION NUMBER: US/10/303,134
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-134-11

Alignment Scores:
Pred. No.: 2,58e-152 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-134-11 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGATGTAATCAATCAATTTTATTGGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyr 60
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTCGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTTAATCCAAATCAAGAAATGAGACCGAACTAATTAIGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAGAACTTTTACGATTATTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTCCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAAAACAGAAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACATATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACCCTTAAAAATAAACTATATTGCTTATGTCTTACACAGTCTTTTAGCAAAATTTT 660

; Sequence 11, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-134-11

Alignment Scores:
Pred. No.: 2,58e-152 Length: 873
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-734-719-12 (1-291) x US-10-303-134-11 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGATGTAATCAATCAATTTTATTGGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyr 60
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTCGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTTAATCCAAATCAAGAAATGAGACCGAACTAATTAIGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAGAACTTTTACGATTATTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTCCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAAAACAGAAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACATATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACCCTTAAAAATAAACTATATTGCTTATGTCTTACACAGTCTTTTAGCAAAATTTT 660
```





Qy	1	MetLysIysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAATAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAAAGAAATTCATTATTCCAAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAAATGAATTTTGATGTATTTAGATGTAACTCAATTTTTATTTTTGAAGATAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr	60
Db	121	CTTGGTAATAAATGCAGAAGCAGTGTTTTACACCCCTGGTTTCTCTCTTGAGCAAATACTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAACAATTTAATCCAAAATCAAGAATAATGAGCCGAACATAATTATGTGTCTTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTGAAAACATTTTTCACGATTATTTTCCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTTTAAACAACATTAAGAATTTAAATGCTTTATTTTAA	360
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysalavalAla	140
Db	361	TTTTCACGAAATTTATTTTCAATCAAGAATAATCCTCAGGGGCTCATATGTGTGCAGTAGGC	420
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATCAAGAANAATTTATCTTCGGGAATTCATTTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaproAspPheLysAsn	180
Db	481	TCCTTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAACTAGCCCTCATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCGCTCACATATATCGGCACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAACCTCAAAANATAAACTATATTGCTTATGTCCTAACAGCTCTTTTACAAAATTTT	660
Qy	221	IleGluLeuAlaproAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysasnIleAsnPhe	260
Db	721	AAAGATATACATCACTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAATAAANAATTAAGAAAAATGTTTATTACAAGTTGATATAAGATCTATTAAAGATTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGATATAAGCATTTATTTCAAGGAAAA	873



FEATURE: CDS  
LOCATION: (1) (873)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-846-219-11

Alignment Scores:  
Pred. No.: 2 58e-152 Length: 873  
Score: 1563.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 9 Gaps: 0

US-10-734-719-12 (1-291) x US-10-846-219-11 (1-873)

Qy 1 MetLysVallelleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
Db 121 CTGGTAAAAAATGCAAGCAGTGTTTACACCCCTGGTTCTTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAGAAATGAGACCACTAATATGTTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTCTAATAAATTTTACGATTTATTTCTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheGluGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTTAATGCTTTATTTAA 360  
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCATATGTTGTCAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATCTACAAATAAATATATATGCTATGTTCTTAACAGTCTTTTAGCAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAATAACTACACT 720  
Qy 241 LysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATATCTCATACCTCTTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780  
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATTAATAATTAAGAAATGTTTATACAGTTTGATATAAAGATCTATTAAGATTA 840

Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAAGCATTATTTCAAAGAAAA 873

# RESULT 12

US-10-821-604-11  
Sequence 11, Application US/10821604  
Publication No. US20040229263A1  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
TITLE OF INVENTION: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/821,604  
CURRENT FILING DATE: 2004-04-08  
PRIOR APPLICATION NUMBER: 10/303,128  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 873  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) (873)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-821-604-11

Alignment Scores:  
Pred. No.: 2 58e-152 Length: 873  
Score: 1563.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 9 Gaps: 0

US-10-734-719-12 (1-291) x US-10-821-604-11 (1-873)

Qy 1 MetLysLysVallelleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
Db 121 CTGGTAAAAAATGCAAGCAGTGTTTACACCCCTGGTTCTTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAGAAATGAGACCACTAATATGTTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTCTAATAAATTTTACGATTTATTTCTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheGluGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTTAATGCTTTATTTAA 360  
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140



Db 361 TTTACGAAATTATTTCAATCAAGAATTACTCTCAGGGTCTATATGTGTGCTAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCCTCACACTATATCGACATAGTAAATAACAGATATAAAGCTTTAGATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATCTCAAAAATAAAACTATATGCTTATGCTTCAACAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGGCCCAATTTTAAATTCAAATTTTATCATACAGAAAAATAACACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAATTTTCAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTAATTTCAAGGAAAA 873

## RESULT 13

US-10-847-983-11  
; Sequence 11, Application US/10847983  
; Publication No. US20040229272A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/847,983  
; CURRENT FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-847-983-11

## Alignment Scores:

Pred. No.:	2,58e-152	Length:	873
Score:	1563.00	Matches:	291
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0

DB: 9 Caps: 0  
US-10-734-719-12 (1-291) x US-10-847-983-11 (1-873)  
QY 1 MetLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATCAAAAAAGTTATTATGCTGAAATGGACCAAGTTTAAAAAGAAATATGATTATTCAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTTGTATGTTTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTTCTTTGAGCAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAAAACATTTTAATCCAAAATCAAGAATATGAGACCGAACTAATATATGTGTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTAATTTTCCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACATTAAAGAAATTAATGCTTATTTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTATTTTCAATCAAGAATTTACCTCAGGGCTCTATATGTGTGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCCGGAAATTCATTTTATCAAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATCTCAAAAATAAAACTATATGCTTATGCTTCTCCTAACAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGGCCCAATTTTAAATTCAAATTTTATCATACAGAAAAATAACACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTAATTTCAAGGAAAA 873

## RESULT 14

US-10-821-573-11  
; Sequence 11, Application US/10821573  
; Publication No. US20040229313A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/821,573  
 ; CURRENT FILING DATE: 2004-04-08  
 ; PRIOR APPLICATION NUMBER: 10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 11  
 ; LENGTH: 873  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(873)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:4  
 US-10-821-573-11  
 Alignment Scores:  
 Pred. No.: 2,586-152 Length: 873  
 Score: 1563.00 Matches: 291  
 Percent Similarity: 100.0% Conservatives: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 Gaps: 9  
 US-10-734-719-12 (1-291) x US-10-821-573-11 (1-873)  
 Qy 1 MetLysValIlellealaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 Db 1 ATGAAAAAGTTATTATTCTCGAAATGACCAAGCTTTAAAAAGAAATGATTATTCAAGG 60  
 Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
 Db 61 CTACCAAAATGATTTTGATGTATTATAGATGTAATCAATTTATTTTGAAGATAAATACATAT 120  
 Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
 Db 121 CTYGTAAAAATGCAAGACAGTGTTTACACCCCTGGTCTCTTTGAGCAATACCTAC 180  
 Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
 Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACAATATATGTGTCTAAT 240  
 Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100  
 Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTCCCT 300  
 Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
 Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValala 140  
 Db 361 TTTTACGAAATTTATTTCAATCAAGAAATATACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
 Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
 Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTTGATTTTAAAAAT 540  
 Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

```
Db      121 CTTGGTAAAAATGCAACAGCTGTTTACACCCCTGGTTTCTCTTTGAGCAATACTAC 180
QY      61 ThrLeuLysHisLeuIleGlnAsnGlnGluThrGluLeuIleMetCysSerAsn 80
Db      181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATATATGTTCTAAT 240
QY      81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAAAACTTTTTACGATTATTTTCCT 300
QY      101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAA 360
QY      121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTGTGAGTAGCC 420
QY      141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY      161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db      481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTCGATTTTAAAAAT 540
QY      181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY      201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTATGCTTAAACAGTCTTTTAGCAATTT 660
QY      221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db      661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT 720
QY      241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721 AAAGATATACTCATACCTTCTAGTAGGCTTTATGGAATTTTCAAAAAATATTAATTTT 780
QY      261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAACTTGATAAAGATCTATTATAGATTA 840
QY      281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841 CCTAGTATATAAGCATTTATTTCAAAAGGAAAA 873
```

Search completed: August 21, 2006, 16:48:35

Job time : 1390 secs

```
GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:      August 21, 2006, 15:36:45 ; Search time 195.5 Seconds
           (without alignments)
           3569.177 Million cell updates/sec

Title:      US-10-734-719-12
Perfect score: 1563
Sequence:    1 MKKVIIAGNPSLKEIDYSR.....KLIKDLLRLPSDIKHVFKGK 291

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext 0.5
               Ygapop 10.0 , Ygapext 0.5
               Fgapop 6.0 , Fgapext 7.0
               Delop 6.0 , Delext 7.0

Searched:    2226170 seqs, 799283156 residues

Total number of hits satisfying Chosen parameters:      4452340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Command line parameters:
-Model=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB.spool/US10734719/runat.21082006.141832.12628/app_query.fasta.1
-DB=Published Applications NA.New -QMT=fastap -SUFFIX=P2N.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NOR=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abes04
-USER=US10734719 @CGN 1.1.286 @runat.21082006.141832.12628 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6 -FGAEXT=7
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DEEXT=7

Database : Published Applications NA.New:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US12_NEW_PUB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US13_NEW_PUB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US14_NEW_PUB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US15_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result  % Query Match Length DB ID Description
No.
1 1563 100.0 873 6 US-10-734-719-11 Sequence 11, Appl
2 1557 99.6 876 6 US-10-734-719-8 Sequence 8, Appl
3 1553 99.4 873 6 US-10-734-719-13 Sequence 13, Appl
4 1534 98.1 876 6 US-10-734-719-4 Sequence 4, Appl
5 1524 97.5 876 6 US-10-734-719-2 Sequence 2, Appl
6 1524 97.5 11474 6 US-10-734-719-1 Sequence 1, Appl
7 1513 96.8 876 6 US-10-734-719-6 Sequence 6, Appl
```

## ALIGNMENTS

## RESULT 1

```
US-10-734-719-11
; Sequence 11, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10734.719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/119,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-734-719-11
```

Sequence 3183, Ap  
Sequence 3113, Ap  
Sequence 1, Appl  
Sequence 1901, Ap  
Sequence 65, Appl  
Sequence 6121, Ap  
Sequence 41, Appl  
Sequence 6973, Ap  
Sequence 8, Appl  
Sequence 17721, A  
Sequence 14292, A  
Sequence 3504, Ap  
Sequence 12295, A  
Sequence 1426, A  
Sequence 10401, A  
Sequence 16, Appl  
Sequence 3347, Ap  
Sequence 188441, A  
Sequence 194699, A  
Sequence 22355, A  
Sequence 29680, A  
Sequence 46, Appl  
Sequence 5949, Ap  
Sequence 80377, A  
Sequence 13344, A  
Sequence 53, Appl  
Sequence 15, Appl  
Sequence 11735, A  
Sequence 18, Appl  
Sequence 282483, A  
Sequence 309123, A  
Sequence 391890, A  
Sequence 482608, A  
Sequence 17359, A  
Sequence 5411, Ap  
Sequence 4623, Ap  
Sequence 12070, A  
Sequence 707, App

Alignment Scores:  
Pred. No.: 3,09e-156 Length: 873  
Score: 1563.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 6 Gaps: 0

US-10-734-719-12 (1-291) x US-10-734-719-11 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
|||  
DB 1 ATGAAAAAGTTATTATTCTGCGAAATGACCAAGTTTAAAGAAATTGATTATTCAAGG 60  
|||  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
|||  
DB 61 CTACCAAAATGATTTCGATGTAATTAGATGAATCAATTTTATTGTAAGATAAATACTAT 120  
|||  
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
|||  
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTGGTTCTCTTTGAGCAATACTAC 180  
|||  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
|||  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATTATGTGTTCTAAT 240  
|||  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100  
|||  
DB 241 TACCAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAACTTTTACGATTATTTTCCT 300  
|||  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
|||  
DB 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTAATGCTTATTTTAA 360  
|||  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
|||  
DB 361 TTTTCAGAAATTTAATTCATCAAGAAATTAATCTCAGGGGTCTATATGTGTCAGTAGCC 420  
|||  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
|||  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480  
|||  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
|||  
DB 481 TCTTATGCTTTTGGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
|||  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
|||  
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
|||  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
|||  
DB 601 GAAAAAACTTACAAAATAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAATTTT 660  
|||  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
|||  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
|||  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
|||  
DB 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTT 780  
|||  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
|||  
DB 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAGTTGATTAAGATCTATTAAGATTA 840  
|||  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
|||  
DB 841 CCTAGTATATAAAGCATTTATTTCAAGGAAAA 873  
|||

RESULT 2

US-10-734-719-8

; Sequence 8, Application US/10734719  
; Publication No. US20060166317A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/734,719  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US 09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CatII) from C. jejuni O:19  
US-10-734-719-8

Alignment Scores:  
Pred. No.: 1,34e-155 Length: 876  
Score: 1557.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.6% Indels: 0  
DB: 6 Gaps: 0

US-10-734-719-12 (1-291) x US-10-734-719-8 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
|||  
DB 1 ATGAAAAAGTTATTATTCTGCGAAATGACCAAGTTTAAAGAAATTGATTATTCAAGG 60  
|||  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
|||  
DB 61 CTACCAAAATGATTTCGATGTAATTAGATGAATCAATTTTATTGTAAGATAAATACTAT 120  
|||  
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyr 60  
|||  
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTTTGAGCAATACTAC 180  
|||  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
|||  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATTATGTGTTCTAAT 240  
|||  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100  
|||  
DB 241 TACCAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAACTTTTACGATTATTTTCCT 300  
|||  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
|||  
DB 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTATTTTAA 360  
|||  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
|||  
DB 361 TTTTCAGAAATTTAATTCATCAAGAAATTAATCTCAGGGGTCTATATGTGTCAGTAGCC 420  
|||  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
|||  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480  
|||  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
|||

Db	481	TC	TATGCTTTTGATACCAACCAAGAAAATCTTTTAAAACTAGCCCTCATTTTAAAAAT	540
Qy	181	Asp	ArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GAT	CGTCGCACATATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA	600
Qy	201	Glu	LysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GA	AAAAACTTACAAAATAAAACTATATTGCTTATGTCCCTAATAGTCTTTTAGCAATTTT	660
Qy	221	Ile	GluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	AT	GAACACTAGGCCCAATTTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT	720
Qy	241	Lys	AspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AA	AGATATACATACCTTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTT	780
Qy	261	Lys	LysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	781	AA	AAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA	840
Qy	281	Pro	SerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CT	AGTATATATAAAGCATTTATTTCAAAGGAAAA	873
RESULT 3				
US-10-734-719-13				
; Sequence 13, Application US/10734719				
; Publication No. US20060166317A1				
; GENERAL INFORMATION:				
; APPLICANT: Gilbert, Michel				
; APPLICANT: Wakarchuk, Warren W.				
; APPLICANT: National Research Council of Canada				
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of				
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics				
; FILE REFERENCE: 019633-000111US				
; CURRENT APPLICATION NUMBER: US/10/734, 719				
; CURRENT FILING DATE: 2003-12-11				
; PRIOR APPLICATION NUMBER: US 09/816,028				
; PRIOR FILING DATE: 2001-03-21				
; PRIOR APPLICATION NUMBER: US 60/118,213				
; PRIOR FILING DATE: 1999-02-01				
; PRIOR APPLICATION NUMBER: US 09/495,406				
; PRIOR FILING DATE: 2000-01-31				
; NUMBER OF SEQ ID NOS: 49				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 13				
; LENGTH: 873				
; TYPE: DNA				
; ORGANISM: Campylobacter jejuni				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (1)..(873)				
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II				
; OTHER INFORMATION: (CstII) from C. jejuni O:36				
US-10-734-719-13				
Alignment Scores:				
Pred. No.:		3,55e-155	Length:	873
Score:		1553.00	Matches:	289
Percent Similarity:		99.3%	Conservative:	0
Best Local Similarity:		99.3%	Mismatches:	2
Query Match:		99.4%	Indels:	0
DB:		6	Gaps:	0
US-10-734-719-12 (1-291) x US-10-734-719-13 (1-873)				
Qy	1	Met	LysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	AT	GAATAAAAGTTATTATTGCTGGAATGCACCAAGTTTAAAAAGAAATTCGATTATTCAAGG	60



```
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTACGATTATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAATTTGGGATGATGATTTTTCACAACTTAAAGATTTTAAAGCTTTATTTAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACAACTAAACAAATAATCTTTTAAATTTGGCTCCTTAATTTAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTCACACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAATAAATAATATATGCTTATGCTTACAGCTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATGAACTAGCGCAATTTAAATTTCAATTTTATCATATACAGAAAAATTAATACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTATTA 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATTAATAATTAAGAAAAATTTTATTTATACAAAGTTGATAAAGATCTATTAAGAT 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 873
```

RESULT 6

```
US-10-734-719-1
; Sequence 1, Application US/10734719
; Publication No. US20060186317A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10734, 719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816, 028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118, 213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495, 406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
```

```
US-10-734-719-1
Alignment Scores:
Pred. No.: 1,13e-150 Length: 11474
Score: 1524.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.5% Indels: 0
Gaps: 0
DB: 6

US-10-734-719-12 (1-291) x US-10-734-719-1 (1-11474)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTCATTATTCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProGlyPhePhePheGluGlnTyrTyr 60
Db 6168 CTTCGTAATAAATGCAAGGCAGTATTTTACAACTCTATTTCTTTTTCGACCAATACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTTAAACATTTAATCCAAATTCAGAAATATGAGACGAACTAATATATGTCTTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATGATGATTTTTCACAACTTAAAGATTTTAAATCTTATTTTAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTTACGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGCGAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATCTATAACAAAAAATCTTTTAAAAATTTGGCTCTCTAATTTAAAA 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTCACAAAAATAAACTATATTTGCTTATGTCTCTAACAGCTTTTAGCAATTTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCAATTTTAAATTTCAAAATTTTATCATAACAGAAAAATAACTACACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATACTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 6827
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAATTAATAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAGATTA 6887
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 6920
RESULT 7
```





US-10-734-719-12 (1-291) x US-11-174-307B-3183 (1-1659)

```
QY 9 AsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPhe 28
Db 377 CACTCGCGCGGTATTCAAGCTGAAGTATGAGAAATGAGAAAGATTCAAGGTTTAT 436
QY 29 -----ArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLys 43
Db 437 ATATATCCAGATGGGATCCCAACAGCTTCTACAGACGCGGAGGAAGCTCACCGGAAG 496
QY 44 LysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyrThrLeuLys 63
Db 497 -----TATGCCAGTGAAGCTATTCTTCAGAAT----- 526
QY 64 HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGln 83
Db 527 -----ATCAGACACAGTGGCTCCGACCGCAA-----AATCCCGATGAG 565
QY 84 AlaHisLeu----- 86
Db 566 GCACACCTCTTCTTCAATCCCATCTCGTGTCAAGATGCGTGGCAAGGCACAGTTAT 625
QY 87 GluAsn-----GluAsnPheValLysThrPheTyrAspTyrPhePro--- 100
Db 626 GAGAATATGACAATAATTGTGCAAAATATGTGGAGAGCTTGATATCCAAATATCCTTAT 685
QY 101 ---AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 686 TGGATAGAACCTTGGGTGTGATCACTTC-----TTTGTC 721
QY 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal----- 134
Db 722 ACTTGTCATGATGTT-----GGTGTGAGGCGCAACAGAGGACTTGAGTTTCTTTGTGAAG 775
QY 135 -----TyrMetCysAlaValAlaIleAlaLeuGlyTyr-----LysGlu 147
Db 776 AATTCCATTTCGACGGGTGCTCCCCAGCTATGATGTGGATTCATTCCACATAAGAT 835
QY 148 IleTyrLeu-----SerGlyIleAspPheTyrGln 157
Db 836 GTGCTCTCCCTCAAGTGTCTACAGCCATTGCTCTTCTGCTGCTGGGAATGATATAGAA 895
QY 158 AsnGlySerSerTyrAlaPhe-----AspThrLys----- 167
Db 896 AACAGAGCTACTCTTGATTTCTGGGTGTGCTCGGTGTCATCGGAACCTCTAAAATCAGAGTTATCTTT 955
QY 168 -----GlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArg---Ser 183
Db 956 GCACGTGTGTGGGAAATGACACAGACTT-----GATATTTCAAACAATAGATTAGT 1009
QY 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr 203
Db 1010 AGGGCTACTGGGCAT-----CTAGTGTACCAGAAAGATTTT 1045
QY 204 TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223
Db 1046 TACAGGAAGAAG---TTTTGTATATGTCTCGGTGGATCAAGGTTAATAGTGTGCGGATA 1102
QY 224 AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyr----- 239
Db 1103 GCAGACTCTATCCATTATGGGTGTATTCCTGTGATATTGCAATTTACTATGACCTTCCT 1162
QY 240 ThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsn 259
Db 1163 TTCAATGATATCTT-----GACTGGGAACAAATTTGCC----- 1195
QY 260 PheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArg 279
Db 1196 -----GTTGTACTCAGGAGAGTGATGTTTACAGCTT---AAACAATATCCTCAAA 1243
QY 280 LeuProSerAsp 283
Db 1244 AATATATCAGAT 1255
```

## RESULT 9

US-11-216-545-3113  
; Sequence 3113, Application US/11216545  
; Publication No. US20060135758A1

## ; GENERAL INFORMATION:

; APPLICANT: MONSANTO Technology, LLC  
; APPLICANT: McLaird, Paul L  
; APPLICANT: Tao, Nengbing  
; APPLICANT: Wu, Kunsheng

; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping  
; FILE OF INVENTION: Soybeans.

; FILE REFERENCE: 38-21 (53659)B

; CURRENT APPLICATION NUMBER: US/11/216,545

; CURRENT FILING DATE: 2005-08-31

; PRIOR APPLICATION NUMBER: US 60/606,062

; PRIOR FILING DATE: 2004-08-31

; NUMBER OF SEQ ID NOS: 8783

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 3113

; LENGTH: 1614

; TYPE: DNA

; ORGANISM: Glycine max

US-11-216-545-3113

## Alignment Scores:

Pred. No.:	0.0294	Length:	1614
Score:	113.00	Matches:	73
Percent Similarity:	39.0%	Conservative:	61
Best Local Similarity:	21.2%	Mismatches:	90
Query Match:	7.2%	Indels:	120
DB:	8	Gaps:	22

US-10-734-719-12 (1-291) x US-11-216-545-3113 (1-1614)

```
QY 9 AsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPhe 28
Db 352 CACTCGCGCGGTTCAGACTGAAGTATGAGAAATGAGAAAGATTCAAGGTTTAT 411
QY 29 -----ArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLys 43
Db 412 ATATATCCAGATGGGATCTTAACACGTTCTTACAGACGCGGAGGAAGCTCACCGGAAG 471
QY 44 LysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyrThrLeuLys 63
Db 472 -----TACGCCAGCGAAGGCTATTCTTCCAGAAC----- 501
QY 64 HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGln 83
Db 502 -----ATCAGAGAGATCGCTTCTGCACCGAA-----AATCCCGATGAG 540
QY 84 AlaHisLeu----- 86
Db 541 GCACACCTCTTCTTCAATTCCTCCATCTCGTGTCAAGATGCGTGGCAAGGCACAGTTAT 600
QY 87 GluAsn-----GluAsnPheValLysThrPheTyrAspTyrPhePro--- 100
Db 601 GAGAATATGACGATAATTGTACAAAACCTATGTGAGAGCTTGATATCCAAATATCCTTAT 660
QY 101 ---AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 661 TGAACAGAACCTTGGGTGCTGATCACTTC-----TTTGTC 696
QY 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal----- 134
Db 697 ACTTGTCATGATGTT-----GGCGTGAAGGCAACAGAAAGGACTTGAGTTTCTTGTGAAG 750
QY 135 -----TyrMetCysAlaValAlaIleAlaLeuGlyTyr-----LysGlu 147
Db 751 AATTCCATTTCAGCTGTGTCTCCCGCAGCTATGATGTGGATTTCATCCCGCATAAAGAT 810
QY 148 IleTyrLeu-----SerGlyIleAspPheTyrGln 157
```

```
Db      811 GTTGCTCTCCCTCAAGTGTCTACAGCCATTGCTCTTCTGCTGGGGGAATGATATAGAA 870
QY      158 AnGlySerSerTyrAlaPhe-----AspThrLys-----
Db      871 AACAGAACTACTCTCTGGATTTTGGGCGGTGCATCGTAACCTCTAAAATCAGAGTTATTCTT 930
QY      168 -----GlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArg---Ser 183
Db      931 GCACGTGTGTGGGAAAATGACACAGAACTT-----GATATTCAAACAATAGAAATTAGT 984
QY      184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr 203
Db      985 AGGGCTACTGGGCAT-----CTAGTGTAACCAAGAGAGATT 1020
QY      204 TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheileGluLeu 223
Db      1021 TACAGAGAGTAAG---TTTGTATATGCTCTGCTGGATCAAGGTTAATAGTCTCGGATA 1077
QY      224 AlaProAsnLeuAsnSerAsnPheilelleGlnGluLysAsnAsnTyr----- 239
Db      1078 GCAGACTCTATCCATTATGGGTGTATTCTGTGATATTGCTTAATTACTATGACCTTCCT 1137
QY      240 ThrLysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsn 259
Db      1138 TTCAATGATATTCTT-----GACTGGAAACAATTGCT----- 1170
QY      260 PheLysLysIleLysLeuAsnValTyrTyrLysLeuLysAspLeuLeuArg 279
Db      1171 -----GTTGTACTCAAGGAGAGTGATGTATACCAGCTT---AAACAAATCCTCAAA 1218
QY      280 LeuProSerAsp 283
Db      1219 AATATATCAGAT 1230

RESULT 10
US-10-734-719-1/c
; Sequence 1, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/734,719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-734-719-1
```

## Alignment Scores:

Pred. No.:	0.461	Length:	11474
Score:	112.00	Matches:	76
Percent Similarity:	40.8%	Conservative:	43
Best Local Similarity:	26.0%	Mismatches:	100
Query Match:	7.2%	Indels:	74
DB:	6	Gaps:	16

US-10-734-719-12 (1-291) x US-10-734-719-1 (1-11474)

```
QY      43 LysLysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyrThrLeu 62
Db      6212 AAAAGAATAGAGATTGTAAATACTCCCTTGCACTTTTACCAGATAGTATTATCTTC 6153
QY      63 Lys---HisLeuileGlnAsnGlnGluTyrGluThrGluLeuileMetCysSerAsn--- 80
Db      6152 AAAATAAAATTTGATTACATCTAAATA-CATCAAAATCATTTGGTAGTCTTGAATAATCAA 6094
QY      81 -----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPheValLysThrPheTyrAspTy 98
Db      6093 TTTCTTTTAAACTTGCTCCATTTCAGCAATAATAAATCTTT-----TTTCATGATTA 6043
QY      98 rPheProAspAlaHisLeuGlyTyrAspPhePheLys----- 110
Db      6042 TTTTCATCTCAAGATTATTATTTTAAAGATTATCAAAATAGATTTTATGTATAATTTAA 5983
QY      111 -GlnLeuLysGluPheAsnAlaTyr-PheLysPheHisGluIleTyr-----P 126
Db      5982 TTATAATAAAATAAGGAATAAATATGTTTAAATTTTCAATCATCTTACCAACTTATAATG 5923
QY      126 heAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrL 146
Db      5922 TCGAACATAATATATAGCAAGGCAATAGAAAGCTGT-----ATCAATCAGACTTTTA 5872
QY      146 ysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspT 166
Db      5871 AAGATATAGAAATAATTGTAGTTGAT----- 5842
QY      166 hrLysGlnGluAsnLeuLysLeuAlaProAspPhe---LysAsnAspArgSerHisT 185
Db      5841 GTCGAAATGATAATAGTATAAATAATATAGCCAAAGAACTACTCTAAAAAGACAAAAGATAA 5782
QY      185 yrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrL 205
Db      5781 AAATAATCCACAATGAA-----AAAAAACTTAGGTCTTTTAAAGACGCAAGATATG 5734
QY      205 ys---IleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheileGluLeuA 224
Db      5733 AAGGTGTGAAGTA-----GCAAACTCTCTTATATAATGTTTTTTAGATCCTG 5686
QY      224 laProAsnLeuAsnSerAsnPhe-----IleIleGlnGluLysAsnA 238
Db      5685 ATGATTATTTGGAACTAAATGCTTTGGAAGAGGTATATAAAATTTTAGATGAACAGGATG 5626
QY      238 sn-----TyrThrLysAspIleLeuileProSerSerGluAlaTyrGlyLysP 254
Db      5625 AAGTTGATTTAGTGTGTTTTTCAATGCTATTGTTGAAAAGTAATGTTATTTTCATATAAAAAGT 5566
QY      254 heSerLysAsnIleAsnPhe-----LysLysIleLysIleLysG 267
Db      5565 TTGACTTTTAATCTCGTGTGTTTATAGCAAAAAGAGTTTGTAAAAAAATTTATTGCCAAGA 5506
QY      267 luAsnValTyrTyr-----LysLeuileLysAsp----- 276
Db      5505 AAAATTATATGGAATGCTGCGGGAACCTTATAAGAAAGAAATTCGTTATTAGAGCTT 5446
QY      277 -----LeuLeuArgLeuProSerAspIleLys 285
Db      5445 TTGCGAGTTTAAGACTCGAGAAAGATGTTAAA 5414
```

## RESULT 11

```
US-10-471-571A-1901
; Sequence 1901, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
```



```
Db 466 GAGATTTCACGCTTCTCTG----- 486
QY 213 ProAsnSerLeuLeuAlaAsnPhelleGluLeuAlaProAsnLeuAsnSerAsnPhelle 232
Db 487 TCTAATAACATGTTGTAACAGATTAATAAGACTCAGCCTTAATCTGAACAGCCAGCTAAA 546
QY 233 IleGlnGluLys---AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr 251
Db 547 GTGATGGAAAAGTTATCAACTTTCGCCACTTCTGGCTCAATGCCATATCAGCATGCTTAT 606
QY 252 GlyLysPheSerLysAsnIle-AsnPhelLysLysIleLysLysGluAsnValTyr 270
Db 607 GAAAGTGTAAGATTATTCATTAACAAGCTCCTGTGTGTTGAACAATCCAGTACAT 664

RESULT 13
US-11-217-529-6121
; Sequence 6121, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6121
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-6121

Alignment Scores:
Pred. No.: 0.494 Length: 2175
Score: 103.00 Matches: 72
Percent Similarity: 36.2% Conservatives: 42
Best Local Similarity: 22.9% Mismatches: 117
Query Match: 6.6% Indels: 84
DB: 8 Gaps: 15

US-10-734-719-12 (1-291) x US-11-217-529-6121 (1-2175)
QY 6 IleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeuPro----- 22
Db 802 CTGGTTCAAAGGGGCCCAATTTTAAAGAAGTC-----ATTCCACTTGAAAC 849
QY 23 AsnAspPheAspValPhe-----ArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 850 GAAGACCTAAAATAATAATAAGAAAAATGTTTCGATTGCGATTGTTTAAAGATGGTGA 909
QY 41 LeuGlyLys----- 43
Db 910 CTAGTACGGTTCTGGACGATCATACTTTAATTAATCTCGGAATCGTAATGGATTTA 969
QY 44 ---LysCysLysAlaValPheTyrThrProGlyPhePheGluGlnTyrTyrThrLeu 62
Db 970 GAAACATGATCATCGACTTCTTCAAGTGGGACTTCTTAGATAGGCTGATAGACTT 1029
QY 63 LysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsn 82
Db 1030 -----TATGATACTGATATCATTCAGATAGCTCC----- 1059
QY 83 GlnAlaHisLeuGluAsnGluAsnPhelValLysThrPheTyrAspTyrPhePro----- 100
Db 1060 -----GTAGAGAGAGAAAAGTTTCGTACAAAAACGTAAGACGGTATCAGATTACTC 1110
```

```
QY 101 -----AspAlaHisLeuGlyTyrAspPhePheLys 110
Db 1111 CAACAATGTGTTCAAAATGCCATTATTAATTAGACGGCATGACCGTTCCCAAGTTTATAAG 1170
QY 111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130
Db 1171 ACATTGTGTCGA--AAGGGTCTATTCAAAGTCTTAGATTATGCAATTTCACATGGAAC 1227
QY 131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150
Db 1228 GATAGTAACGTTTCGCATTTTGGCTACCGATCTATCATTCATCCATCATCGAGCATGACATA 1287
QY 151 SerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsn 170
Db 1288 TTGCTGATTCACACGTTCAAAATGAANAATCTTTC-----AAACAACAACAC 1335
QY 171 LeuLeuLysLeuAlaProAspPheLys---AsnAspArgSerHisTyrIleGlyHisSer 189
Db 1336 -----AAATTGGCCCCGGATGACAAAATGTTCTACACGAAGCAATTCACAGGATTACAAC 1389
QY 190 LysAsnThrAspIleLysAlaLeuGluPheLeu----- 200
Db 1390 TCCAGCAGTCACTCCAAGCTATTATTATAATCTTTCAACGATTCTTTAAACCGCAAAAAGT 1449
QY 201 -----GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu 216
Db 1450 CCGGGTTACGGCAACAAGTTGTGCAAGCGCTAAACACGTTGTTTCAACCGCAGAGGATGT 1509
QY 217 LeuAlaAsnPhelleGluLeuAlaProAsnLeuAsnSerAsnPhelleIleGlnGluLys 236
Db 1510 CTCAGTAAT-----GGTGAAGGTTTCGTATCATATCATGATAGTAGATCC 1551
QY 237 AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLys 256
Db 1552 AACACGAATCAAAAAT-----ACCTCTGAAGACTTCCCAAAATTCGGTTAT 1599
QY 257 AsnIleAsnPhelLysLysIleLysLysGluAsnValTyrTyr 271
Db 1600 GGATTGAATTCGATTCAATTAATTG---GACAAATACAACTAC 1641

RESULT 14
US-11-021-837-41
; Sequence 41, Application US/11021837
; Publication No. US20060140972A1
; GENERAL INFORMATION:
; APPLICANT: Alm, Richard
; APPLICANT: Manning, Paul
; APPLICANT: McLaughlin, Robert
; APPLICANT: McCormack, Kathleen
; TITLE OF INVENTION: Staphylococcus Saprophyticus Nucleic Acids and Polypeptides
; FILE REFERENCE: 101324-US
; CURRENT APPLICATION NUMBER: US/11/021,837
; CURRENT FILING DATE: 2004-12-23
; PRIOR APPLICATION NUMBER: US 60/533534
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/600680
; PRIOR FILING DATE: 2003-08-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 41
; LENGTH: 109669
; TYPE: DNA
; ORGANISM: Staphylococcus saprophyticus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (46850)..(46850)
; OTHER INFORMATION: n is a, c, g, or t
US-11-021-837-41

Alignment Scores:
Pred. No.: 107 Length: 109669
Score: 101.50 Matches: 55
```



```

Db 688 ---TTGAAAAATTATCATCTT---GATACGCTGTAAACGGAAAAATCAAATAAATTTAATT 741
QY 224 -----AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsn----- 237
Db 742 GTTAATTTTGCAGTTAATCTTTCACAAAGTAATGTCATCAAAAATAGTAGATTTCACGAAC 801
QY 238 -----AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255
Db 802 ACCTTGAATAATTTGAAGACAAACATGTT-----TCGAAAGCAGGATCTAAATTTAAA 855
QY 256 Lys---AsnIleAsnPhelLysLysIleLysGlu----- 267
Db 856 AATATCAATGTCATTTTAAATGCTAATAAGCTGTTGAATCAGGTAAGGATTTTGTAGCT 915
QY 268 AsnValTyrTyrLysLeuIle 274
Db 916 AATATTGGCAACAAATGTT 936

```

Search completed: August 21, 2006, 16:05:44  
Job time : 218.5 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 14:42:06 ; Search time 4772.5 Seconds

(without alignments)  
5848.727 Million cell updates/sec

Title: US-10-734-719-14

Perfect score: 1564

Sequence: 1 MKKVIIAGNPSLKEIDYSR.....KLIKDLRLPSDIKHVFKGK 291

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB.spool/US10734719/runat 21082006 141810 12316/app\_query.fasta\_1  
-DB=GenEmbl -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05p  
-USR=US10734719 @CGN 1 1 8663 @runat 21082006 141810 12316 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_env.\*  
2: gb\_pat.\*  
3: gb\_ph.\*  
4: gb\_pl.\*  
5: gb\_pr.\*  
6: gb\_ro.\*  
7: gb\_sts.\*  
8: gb\_sv.\*  
9: gb\_un.\*  
10: gb\_vi.\*  
11: gb\_ov.\*  
12: gb\_htg.\*  
13: gb\_in.\*  
14: gb\_om.\*  
15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	100.0	873	2	AX934436 Sequence
2	1564	100.0	873	2	AR481787 Sequence
3	1564	100.0	873	2	AR527386 Sequence

4	1564	100.0	873	2	AR609665	Sequence
5	1564	100.0	873	2	AR689943	Sequence
6	1564	100.0	873	2	AR691839	Sequence
7	1564	100.0	12390	15	AF401528	Campyloba
8	1559	99.7	876	2	AX934431	Sequence
9	1559	99.7	876	2	BD249794	Campyloba
10	1559	99.7	876	2	AR271703	Sequence
11	1559	99.7	876	2	AR481785	Sequence
12	1559	99.7	876	2	AR527384	Sequence
13	1559	99.7	876	2	AR609663	Sequence
14	1559	99.7	876	2	AR689941	Sequence
15	1559	99.7	876	2	AR691837	Sequence
16	1559	99.7	4749	15	AF305571	Campyloba
17	1559	99.7	6047	15	AY661458	Campyloba
18	1559	99.7	24425	15	AY422197	Campyloba
19	1559	99.7	24437	15	AF167344	Campyloba
20	1553	99.3	873	2	AX934434	Sequence
21	1553	99.3	873	2	AR481786	Sequence
22	1553	99.3	873	2	AR527385	Sequence
23	1553	99.3	873	2	AR609664	Sequence
24	1553	99.3	873	2	AR689942	Sequence
25	1553	99.3	873	2	AR691838	Sequence
26	1553	99.3	11474	15	AF215659	Campyloba
27	1524	97.4	876	2	AX934427	Sequence
28	1524	97.4	876	2	BD249792	Campyloba
29	1524	97.4	876	2	AR271701	Sequence
30	1524	97.4	876	2	AR481783	Sequence
31	1524	97.4	876	2	AR527382	Sequence
32	1524	97.4	876	2	AR609661	Sequence
33	1524	97.4	876	2	AR689939	Sequence
34	1524	97.4	876	2	AR691835	Sequence
35	1521	97.3	876	2	AX934425	Sequence
36	1521	97.3	876	2	BD249791	Campyloba
37	1521	97.3	876	2	AR271700	Sequence
38	1521	97.3	876	2	AR481782	Sequence
39	1521	97.3	876	2	AR527381	Sequence
40	1521	97.3	876	2	AR609660	Sequence
41	1521	97.3	876	2	AR689938	Sequence
42	1521	97.3	876	2	AR691834	Sequence
43	1521	97.3	876	15	AF216647	Campyloba
44	1521	97.3	6047	15	AY644679	Campyloba
45	1521	97.3	11442	15	AF400048	Campyloba

#### ALIGNMENTS

RESULT 1	AX934436	Sequence 13 from Patent WO02074942.	873 bp	DNA	linear	PAT 05-JAN-2004
LOCUS	AX934436					
DEFINITION	AX934436					
ACCESSION	AX934436					
VERSION	AX934436.1	GI:40641682				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
CDS						

Campylobacter jejuni  
 Campylobacter jejuni  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 Campylobacteraceae; Campylobacter.  
 Gilbert M. and Wakarchuk, W.  
 Campylobacter glycosyltransferases for biosynthesis of gangliosides  
 and ganglioside mimics  
 Patent: WO 02074942-A 13 26-SEP-2002;  
 National Research Council of Canada (CA)  
 Location/Qualifiers  
 1..873  
 /organism="Campylobacter jejuni"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:197"  
 1..873  
 /note="unnamed protein product; Campylobacter  
 alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
 jejuni O:36"  
 /codon\_start=1



```
/transl table=11
/protein_id="CA504171.1"
/db_xref="GI:40841683"
/translations="MKKVIAGNGPSLKEIDYSRLPNDFVFCNCFYFEDKYYLGGK
CKTVFTYPPFFQYTYTLKHLIQNOYETELIMCSNYNQAHLENENFVKTFYDFPDA
HLGYDFKOLKEFNAYFKHEIYFNQRI TSGVYMCVAVALGKYEIYLSGIDFYQNGS
SYAFDFTKOENLLKLPDKNDRSHYIGHSKNTDI KALEFLEKTYKIKLYCLCPNSLLA
NFTLEAPNLNSNFIQEKNNYTKDILLPSEAYGKFSKNINFKKIKENVYYKLKD
LLRPSDIKHYPKKG"
```

## ORIGIN

```
Alignment Scores:      5.4e-189      Length:      873
Pred. No.:             1564.00      Matches:      291
Score:                  100.0%      Conservative:  0
Percent Similarity:    100.0%      Mismatches:   0
Best Local Similarity: 100.0%      Indels:       0
Query Match:          100.0%      Gaps:         2
DB:

US-10-734-719-14 (1-291) x AX934436 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTGATTATCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAAAATGATTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAGAGTGTATTTACACCCCTTAATTTCTTTTGGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGCTGCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAAATATATTTGCTTATGCTTAATAGTCTTTTAGCAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTACGCGCAATTTTAAATTTCAAAATTTTATCATCAAGAAAAATAACTACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAAATATTAATTTT 780
```

```
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAAATTAAGAAAAATGTTTATACAGTTGATAAAGATCTATTATAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTAGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 2
AR481787
LOCUS AR481787
DEFINITION Sequence 13 from patent US 6699705.
ACCESSION AR481787
VERSION AR481787.1 GI:47243422
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 13 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAN;
FEATURES
source 1. .873
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:      5.4e-189      Length:      873
Pred. No.:             1564.00      Matches:      291
Score:                  100.0%      Conservative:  0
Percent Similarity:    100.0%      Mismatches:   0
Best Local Similarity: 100.0%      Indels:       0
Query Match:          100.0%      Gaps:         2
DB:

US-10-734-719-14 (1-291) x AR481787 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAGAGTGTATTTACACCCCTTAATTTCTTTTGGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGCTGCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAAATATATTTGCTTATGCTTAATAGTCTTTTAGCAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTACGCGCAATTTTAAATTTCAAAATTTTATCATCAAGAAAAATAACTACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAAATATTAATTTT 780
```

```
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCATTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTCTTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCAAAATAAACTATATGCTTAGTCCTTAATAGTCTTTTAGCAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATCAAGAAAAATAAATACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 3
AR527386
LOCUS AR527386 873 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 13 from patent US 6723545.
ACCESSION AR527386
VERSION AR527386.1 GI:53914407
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,4-GalNAC transferase activity
JOURNAL Patent: US 673545-A 13 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..873 /organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 5.4e-189 Length: 873
Pred. No.: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-14 (1-291) x AR527386 (1-873)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerIleuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTCGCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGAATGATTTAGATGTAATCAATTTTATTTTGAAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
```

```
181 ACTTTAAACATTTAAATCCAAATATCAAGATATAGACCGAACTAAATATATGTGTTCTAAT 240
81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
241 TACAACCCAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATATATTTCTCT 300
101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
301 GATGCTCATTTGGGATGATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
361 TTTCAACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC 420
141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTGATTTTATCAAAATGGGTCA 480
161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTTA 600
201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
601 GAAAAAATTTACAAAATAAACTATATGCTTAGTCCTTAATAGTCTTTTAGCAATTTT 660
221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATCAAGAAAAATAAATACTACTAC 720
241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
281 ProSerAspIleLysHisTyrPheLysGlyLys 291
841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 4
AR609665
LOCUS AR609665 873 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 13 from patent US 6825019.
ACCESSION AR609665
VERSION AR609665.1 GI:56664965
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 13 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..873 /organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 5.4e-189 Length: 873
Pred. No.: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
```

Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-10-734-719-14 (1-291) x AR609665 (1-873)			
QY	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
DB	1	ATGAAAGAAAGTTATTATCTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCACAGG	60
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
DB	61	CTACCAAAATGATTGGATGATTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120
QY	41	LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
DB	121	CTTGTGTAAGAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC	180
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
DB	181	ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAATTAATGTTCTAAT	240
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
DB	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCTA	300
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
DB	301	GATCGCTACACTATATCGACATAGTAAAAATTAATTTTAAACAACTAGCCCTGATT	360
QY	121	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
DB	421	ATAGCCCTAGGATACAAAGAAATTTAATCTTCGGGAATTTGATATGTTGAGTATTTCTA	480
QY	141	LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
DB	441	CTTGTGTAAGAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC	120
QY	161	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
DB	481	ACTTTAAACATTTTAAATCCAAATCAAGAAATTAATCTTCGGGAATTTGATATGTTGAGTATTTCTA	240
QY	181	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
DB	541	GATCGCTACACTATATCGACATAGTAAAAATTAATTTTAAACAACTAGCCCTGATT	300
QY	201	GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe	220
DB	601	GAAAACTTACAAATTAAGAAATGTTTATTAAGATTTGNTAAGATTA	360
QY	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr	240
DB	661	ATAGAACTAGCGCAAAATTTAAATTTTAAACAACTAGCCCTGATT	300
QY	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
DB	721	AAAGATATATCTCATACCTCTAGTAGGCTTATGSAATTTTCAAAAAATATTAATTTT	320
QY	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
DB	781	AAAAAATTAAGAAATGTTTATTAAGATTTGNTAAGATTTAAGATTA	340
QY	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
DB	841	CCTAGTATATAAGCAATTTATTTCAAGGAAAA	873
RESULT 5			
LOCUS	AR689943	873 bp	DNA
DEFINITION	Sequence 13 from patent US 6905867.	linear	PAT 12-SEP-2005
ACCESSION	AR689943		
VERSION	AR689943.1	GI:74471951	
KEYWORDS			

```

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATCACTCTTAGTGAGGCTTAGGAAAATTTTCAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAATAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 6
LOCUS AR691839 873 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 13 from patent US 6911337.
ACCESSION AR691839
VERSION AR691839.1 GI:74479550
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAc transferase
JOURNAL Patent: US 6911337-A 13 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..873
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 5.4e-189 Length: 873
Pred. No.: 1564.00 Matches: 291
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 2 Gaps: 0
DB:

US-10-734-719-14 (1-291) x AR691839 (1-873)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGCTTATTATTGCTGGNAATGGACCAAGTTTAAAAGAAATTTGATTATCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTATTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACAATTAAATCCAAAATCAAGAATATGAGACCGACTAAATATGTCTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnAtgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAAATTTATTTCATCAATCAAGAATTTACCTCAGGGGCTTATATGTGTGCAGTAGCC 420

```

```

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACACTAGCCCCCTGATTTTAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATGCTCACACTATATCGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAACTATATTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnLeuTyrThr 240
Db 661 ATAGAAGCTAGGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAAATAACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATCACTTCTTAGTGAGGCTTATGAAAAATTTTCAAAAATATTTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATATAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 7
LOCUS AF401528 12390 bp DNA linear BCT 03-JAN-2002
DEFINITION Campylobacter jejuni O:36 heptosyltransferase I (waac) gene, partial cds; lipid A biosynthesis acyltransferase (htcB), putative two-domain glycosyltransferase, putative glycosyltransferase; truncated beta-1,4-N-acetylgalactosaminyltransferase (cgta-I), beta-1,3-galactosyltransferase (cgtB), alpha-2,3-sialyltransferase (cst-II), sialic acid synthase (neuB), putative N-acetylglucosamine-6-phosphate 2-epimerase (neuC1), beta-1,4-N-acetylgalactosaminyltransferase (cgta-II), CMP-Neu5Ac synthetase (neuA), putative acetyltransferase, and putative glycosyltransferase (waav) genes, complete cds; and heptosyltransferase II (waaf) gene, partial cds.
ACCESSION AF401528
VERSION AF401528.1 GI:15718482
KEYWORDS
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 12390)
AUTHORS Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E., Michniewicz,J., Cunningham,A.W. and Wakarchuk,W.W.
TITLE The Genetic Bases for the Variation in the Lipid-oligosaccharide of the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF SIALYLATED GALACTOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
JOURNAL J. Biol. Chem. 277 (1), 327-337 (2002)
PUBMED 11689567
REFERENCE 2 (bases 1 to 12390)
AUTHORS Gilbert,M., Michniewicz,J., Karwaski,M.F., Cunningham,A. and Wakarchuk,W.W.
TITLE Direct Submission
JOURNAL Submitted (22-JUL-2001) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada
FEATURES
source Location/Qualifiers
1..12390
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43456"

```

gene	<div>/serotype="O:36" /db_xref="ATCC:43456" /db_xref="taxon:197" &lt;1..348 /gene="waaC" &lt;1..348 /gene="waaC" /codon_start=1 /transl_table=11 /product="heptosyltransferase I" /protein_id="AAL05984.1" /db_xref="GI:15718483" /translation="FAKEVLNLSGIDETHIELAPFNLELMAFQWMDLIIGDSGP THLALANKASITIPGATSYRNAQTTHINKIIDTGKKIQNAKHIDKSFCTRIBEE DIFKLAKGLLNKK" 338..1225 /gene="htrB" 338..1225 /gene="htrB" /codon_start=1 /transl_table=11 /product="lipid A biosynthesis acyltransferase" /protein_id="AAL05985.1" /db_xref="GI:15718484" /translation="MKNSDRIYLSLVYILKFFVTMPDCILHFLALIVARIAFLNKK HRKIINTNLQICFPQYQTERDKLSKIYENFAQGIQCLQONQTTKEKILNKVAFIN ENFLDALAKKPIIFTTHAYGNWEILSLUAYAAKYGAISIVGKKLASEVMYELSQSR TOFDLELDKKGPIQRLMSALKKRALGILTDQDCVENISVRKLKFNKVENYOMGASL IAQRNALIIPVYAYKEGKGPIEFKFKARDSQNASLEELTYQAQSCSEMIKKRPWEY FFHRFASYNIEIYKGA" 1222..2769 /note="orf3" /codon_start=1 /transl_table=11 /product="putative two-domain glycosyltransferase" /protein_id="AAL05986.1" /db_xref="GI:15718485" /translation="WNLKQISVIIIVKNAEQTILLECILSLKDFDEIILLNSESIDNTL NIANEFKDFANLYIYHNAPIFIGALKUNLALSAYKNDWILSIDADEVLENECIKEIKN LELQENITIALSKYKGEWIKACGSPDYVLRIPFNKFTFRNDLNVHESIVLPSNV KHLKNGLUKHYAFRDISHLINKMYSSILMAKQNIHKKSGLVLANRAFMTFFRNYF LKNGFLYQGGTISVCNALGTFQVKMYLQKQKPTKCALIITYNOKERLKLVD SVKNITELPNEVLADDSKEDTARLIEYQKDFPCSLKHIMQEDSGFRAAKRNNAI RVSKSEYIILIDGDMILEKDFIKNHFEPAKRKVILQGSRVILNKSESELKNDYSL AFSKDKFSYKNSFLAKIITYKISKIDQKFFNTKEIILKSKTNMSPFKADFELDGFN ENFIWGRDSDSFVARFLFNKGIFFRLKFAIAYHYHEENSKOMLESNHQIYLDTIK NKKISWR" 2760..3944 /note="orf4" /codon_start=1 /transl_table=11 /product="putative glycosyltransferase" /protein_id="AAL05987.1" /db_xref="GI:15718486" /translation="MEIKHEENRCSVTNIQRRRIPERMENSVINOTYTNLEILVNDG SIDENSLAKETTLKDERFILFDKNGGLSSARNVGIYFSGEYKLNKQTHRIKENS LIEFDQDGNPNYIKAYKSSQAFNEKDLTNTFYPNIDYIIFLSDSDWELNCEBEC VIRMKNDVLDWFDHCDTYEDNIKNHKKTRMEIFDPKKECIITPKEYANRALSVSGRD ISFGWNGMDLDFNFKKIKLKFINFIINEDIHFGIILFASANKIYILSQKLYLCRLRAN SISNHDKITTKANVSFYFKDIYETFGENAKAKNVLKAASRVTTALKLIEFFKDKOSE NALAIKETPLCPYAKKALMIKKFKDPLNKLEQLVLIKPFIQTKLPDYDIWKFWQKIKK IN" 3988..4248 /gene="cgtA-I" 3988..4248 /gene="cgtA-I" /note="orf5-I; premature translation stop due to a one bp deletion" /codon_start=1 /transl_table=11 /product="truncated beta-1,4-N-acetylgalactosaminyltransferase" /protein_id="AAL05988.1"</div>
CDS	
gene	<div>/db_xref="GI:15718487" /translation="MLFQSYFVKIICLFIPIFRKIRHKIKKTFLLKNIQORDKIDSYLPK KILIQINKYNNEDLIKLKAITGGGIKDILIMKNLKIQNLL" complement(5016..5921) /gene="cgtB" complement(5016..5921) /gene="cgtB" /codon_start=1 /transl_table=11 /product="beta-1,3-galactosyltransferase" /protein_id="AAL05989.1" /db_xref="GI:15718488" /translation="MFKISIIPLTYNVEQYIARIESCINQTFKDIEIIVVDDCGNDN SINIAKEYSKDKRIKIIHNEKNLGLLARVEGVKVPANSPYIMFLDPDDYLELNACEE CIKLDOQEDVLAVFENAIIVESNVIYKKTDFNSGFYSKKEFVKIKIATKNLYWTMWG KLIRKKLYLEAFASLRLEKDVKNMAEDVLLYPMLSQAQKIAMNCNLHYHVPNNNS ICNTKNEVLVKNNIQELQVLNLRQNYIILNKYCSVLYVLIKIYLIQIYIKIRTKLM VTLAKINILTKILFKYKFLKQC" 6010..6885 /gene="cst-II" 6010..6885 /gene="cst-II" /codon_start=1 /transl_table=11 /product="alpha-2,3-sialyltransferase" /protein_id="AAL05990.1" /db_xref="GI:15718489" /translation="MKKVIIAGNPSLKEIDYSRLPNDFVFCRQNFYFEDKYLGKK KLTGYTFPNFFEQYTYFKHEIQNOEYETELIMCSNYNQAHLENENFVKTYDYFPDA HGVYDFKQLKFEFNAYKPEHIIQTSVYMCVAIALGYKEIYLSGLIDFYQNGS SYAFDTQENLLKLPADPNDRSHYIGHSKNTDIIKALEFETKYIKLYCLCPNSLLA NETELAPNLNSFIIOEKNNYTKDILIPSEAYGKFSKNINFKIKIKENVVYKLKD LLRLPSDIKHYPKGN" 6886..7926 /gene="neuB1" 6886..7926 /gene="neuB1" /codon_start=1 /transl_table=11 /product="sialic acid synthase" /protein_id="AAL05991.1" /db_xref="GI:15718490" /translation="MKEIKIQNIIISEKAPLVVPEIGINHNGSLELAKIMVDAAFSA GAKIKHOTHIVEDEMSKAAKVIIPGNAKISIVEIMOKCALDYKDELALKEVTEKLG VYLSTPFRAGANRLEDMGVSAPKIGSGECNNYPLIKHIAFPKPMIVSTGMNLSIEI KPIVTKILLONEIPFVLMHTNLXYTPHNLVRLNAMLELKEFSCWGLSDHTIDLAC LGAVALGALCERHFTDSMHRSGPDIVCSMDTKALKELIIQSEQMAIMKGNNSKKA KOEQVTDFAFASVWSIKDIKKGEVLSDMNIWVRPGLGGISAAEFENILGKKALRDI ENDTQLSYEDFA" 7923..9047 /gene="neuC1" 7923..9047 /gene="neuC1" /codon_start=1 /transl_table=11 /product="putative N-acetylglucosamine-6-phosphate 2-epimerase" /protein_id="AAL05992.1" /db_xref="GI:15718491" /translation="MKKILFITGRADYSKIKSLMYRVQNSSEFELYIFATGMHLSKN FGYTVKELYKNGFKNIIYEFINYDKYOTDKALATTIDGFSRYVNELKPDLIWVHGDRI EPILAAIVGALNNILVAHIEGGEISGTDLSLRHAIKLAHILVNDEFKRLMQLG EDEKSIIFGSPDLELNDNKSISEAKKIYDINENTALLMHFPVTEITSIKNQAD NLVKALIQSNKNYIVIVPNNDLGFELIYSEELKNPRFKLFPSLRPEYFITLLKNA DLFIGNSSCILKEALYIKTAGILVGRQNGRLGNENTLKVNASDEILKAINTIHKKQ DLPSTKLEILDSSKLPPLEYLLSGEFPKLSLTKQKFKDIKHA" 9040..9987 /gene="cgtA-II" 9040..9987 /gene="cgtA-II" /note="orf5-II" /codon_start=1 /transl_table=11</div>
CDS	
gene	<div></div>
CDS	

```
/product="beta-1,4-N-acetylglucosaminyltransferase"  
/protein_id="AAU05993.1"  
/db_xref="GI:15718492"  
/translation="MLKKIISLYKRYISISKLVLDNEHFIKENKNYIGKKHGFDFD  
EKQVNSPLNPGWFIIRVNEAULRVSLESLIPALQORGLIANDCDDGSEELILBFC  
KQYPNFIKKYFKVDLENPKNEENKLYSYNWAASFPLDEWFIKIDVDHYDADKL  
YKSFYIDQENKALCYPRINFILNGNIYVQNSGNYFGTGGDQLIKRNSFIERR  
VSKKSOWIDPKGLIEELYSEQVLSQVKILOAPLQWHPALKYRRNDYQQVLDILS  
LREFOAFHRKSKEAKKIDFAMLRKPVIEQILKKFQGEIK"  
9984..10649  
/gene="neuA1"  
CDS  
9984..10649  
/gene="neuA1"  
/codon_start=1  
/transl_table=11
```

## Alignment Scores:

```
Pred. No.: 6,66e-188 Length: 12390  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 15 Gaps: 0
```

US-10-734-719-14 (1-291) x AF401528 (1-12390)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLysAspTyrSerArg 20  
Db 6010 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCACAGG 6069  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 6070 CTACCAAAATGATTTTGATGATTATTAGATGATCAATTTTATTGTAAGATAAATACTAT 6129  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 6130 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 6189  
Qy 61 ThrLeuLysHisLeuLysGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
Db 6190 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGCACTAATTTATGTGTTCTAAT 6249  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100  
Db 6250 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 6309  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 6310 GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTTAATGCTTTATTAA 6369  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 6370 TTTTCAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 6429  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 6430 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 6489  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 6490 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAANAAT 6549  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 6550 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 6609  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 6610 GAAAAAATTACAAAAATAAACTATATTGCTTATGTCCTAATATGCTCTTTTAGCAAAATTTT 6669  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlyLysAsnAsnTyrThr 240  
Db 6670 ATGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATCAAGAAAAATAAATACTACCT 6729
```

```
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 6730 AAAGATATACATCATCTTCTAGTGAGCTTATGCGAAATTTTCAAAAAATTTAATTTT 6789  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 6790 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTTGATAAAAGATCTATTAGATT 6849  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 6850 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 6882
```

## RESULT 8

```
AX934431  
LOCUS AX934431 876 bp DNA linear PAT 05-JAN-2004  
DEFINITION Sequence 8 from Patent WO02074942.  
ACCESSION AX934431  
VERSION AX934431.1 GI:40641678
```

```
KEYWORDS  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.
```

```
REFERENCE  
1 Gilbert M. and Wakarchuk W.W.
```

```
AUTHORS  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics
```

```
JOURNAL Patent: WO 02074942-A 8 26-SEP-2002;  
National Research Council of Canada (CA)
```

```
FEATURES  
Location/Qualifiers  
1..876  
/organism="Campylobacter jejuni"  
/mol_type="unassigned DNA"  
/db_xref="taxon:197"
```

```
CDS  
1..876  
/note="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
jejuni O:19"  
/codon_start=1  
/transl_table=11  
/protein_id="CAF04169.1"  
/db_xref="GI:40641678"  
/translation="MKKVIITAGNGPSLKEIDYSRLPNDPFRCNQVFEDKYYLGGK  
CKAVYTPNFFEQYITLKHILQOEYETELIMCSNYNQAHLENENFVKTFDYFPDA  
HLGYDFPFQKLKEFNAYFKFHEIYFNQRTSGVYMCVAIALGYKIKLYLSGIDFYQNGS  
SYAPDTQENLLKLAPDKNDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLA  
NFIELAPLNSNFIIOEKNNYTKDILIPSEAYGKFSKNINPKIKIKENVYKLIKD  
LLRLPSDIKHVFKCK"
```

## ORIGIN

```
Alignment Scores:  
Pred. No.: 2.34e-188 Length: 876  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
DB: 2 Gaps: 0
```

US-10-734-719-14 (1-291) x AX934431 (1-876)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLysAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCACAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTTGATGATTATTAGATGATCAATTTTATTGAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
```

Qy	61	ThrLeuLysHisLeuLeuGlnAsnGlnGluThrGluLeuLeuMetCysSerAsn	80
Db	181	ACTTTAAACATTTAATCAAAATCAAGAATATGACCCGAACATAATATGTGTTCTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAenPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCAAGCTCATCTAGAAATGAAATTTGTAAACAACTTTTACGATTATTTTCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGATATGATTTTAAACAACTTTTAAAGAAATTAATGCTTATTTTAA	360
Qy	121	PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGCGAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAACTTACAAAAATAAACTATATGCTTATGCTCTATATAGCAAAATTTT	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATATCTATACCTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAAATAAAAAATAAGAAATGTTTATTAACAAGTTGATAAAAGATCATTAATAAGATTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGATATAAGCATTTATTTCAAAAGGAAAA	873
FEATURES	source	Location/Qualifiers	
ORIGIN			
Alignment Scores:			
Pred. No.:	2.34e-188	Length:	876
Score:	1559.00	Matches:	290
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	99.7%	Indels:	0
DB:	2	Gaps:	0
US-10-734-719-14 (1-291) x BD249794 (1-876)			
Qy	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg	20
Db	1	ATGAAAAAGCTTATTTATGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTC	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAATGATTTTGAATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120
Qy	41	LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	121	CTTGTAATAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC	180
Qy	61	ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn	80
Db	181	ACTTTAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATATGTGTTCTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAenPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCAAGCTCATCTAGAAATGAAATTTGTAAACAACTTTTACGATTATTTTCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGATATGATTTTAAACAACTTTTAAAGAAATTTAATGCTTATTTTAA	360
Qy	121	PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGCGAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAACTTACAAAAATAAACTATATGCTTATGCTCTATATAGCAAAATTTT	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATATCTATACCTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAAATAAAAAATAAGAAATGTTTATTAACAAGTTGATAAAAGATCATTAATAAGATTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGATATAAGCATTTATTTCAAAAGGAAAA	873
RESULT	9		
LOCUS	BD249794	876 bp	DNA linear
DEFINITION	Campylobacter glycolsyltransferases for biosynthesis of gangliosides and ganglioside mimics.		
ACCESSION	BD249794		
VERSION	BD249794.1	GI:33059564	
KEYWORDS	JP 2002535992-A/5.		
SOURCE	Campylobacter jejuni		
ORGANISM	Campylobacter jejuni		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.		
AUTHORS	1 (bases 1 to 876)		
TITLE	Gilbert M. and Wakarchuk, W.W.		
JOURNAL	Campylobacter glycolsyltransferases for biosynthesis of gangliosides and ganglioside mimics		
COMMENT	Patent: JP 2002535992-A 5 29-OCT-2002; NATIONAL RESEARCH COUNCIL OF CANADA		
	OS	Campylobacter jejuni	
	PN	JP 2002535992-A/5	
	PD	29-OCT-2002	
	PF	01-FEB-2000	JP 2000597438
	PR	01-FEB-1999	US 60/118213,31-JAN-2000
	PI	06/495406	PI
	PC	C12N15/09.C12N15/21.C12N9/10.C12N9/88.C12N15/00	CC
	CC	C12N15/09.C12N15/21.C12N9/10.C12N9/88.C12N15/00	CC

```
Db 721 AAGATATATCTCATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
Qy 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTAAGAAAATGTTTATCAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAGGAAAA 873

RESULT 10
LOCUS AR271703 876 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 8 from patent US 6503744.
ACCESSION AR271703
VERSION AR271703.1 GI:29703248
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6503744-A 8 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 2.34e-188 Length: 876
Pred. No.: 1559.00 Matches: 290
Score: 99.7% Conservative: 0
Percent Similarity: 99.7% Mismatches: 1
Best Local Similarity: 99.7% Indels: 0
Query Match: 99.7% Gaps: 2
DB:

US-10-734-719-14 (1-291) x AR271703 (1-876)

Qy 1 MetLysLysValLleAlaGlyAsnGlyProSerLeuLysGluLleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAAGTCAAGAGAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuLleGlnAsnGlnGluTyrGluThrGluLeuLleMetCysSerAsn 80
Db 181 ACITTTAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATATTATGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTTAAA 360
Qy 121 PheHisGluLleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLleTyrLeuSerGlyLleAspPheTyrGlnAsnGlySer 160
```

```
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCCTTAGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrLleGlyHisSerLysAsnThrAspLleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACATATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysLleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATATGCTTATGCTCTAATAGTCTTTTATAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheLleLleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720
Qy 241 LysAspLleLeuLleProSerSerGluAlaTyrGlyLysPheSerLysAsnLleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780
Qy 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTTAAAGAAATGTTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAGGAAAA 873

RESULT 11
LOCUS AR481785 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 8 from patent US 6699705.
ACCESSION AR481785
VERSION AR481785.1 GI:47243420
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 8 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 2.34e-188 Length: 876
Pred. No.: 1559.00 Matches: 290
Score: 99.7% Conservative: 0
Percent Similarity: 99.7% Mismatches: 1
Best Local Similarity: 99.7% Indels: 0
Query Match: 99.7% Gaps: 2
DB:

US-10-734-719-14 (1-291) x AR481785 (1-876)

Qy 1 MetLysLysValLleAlaGlyAsnGlyProSerLeuLysGluLleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
```



```
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
    |||||
Db 181 ACTTTTAAAAACATTTAAATCCAAATCAAGAAATATGAGACCGAACTAATATATGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAGAACTTTTACGATTAATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTCATATGTTGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAAGATCTATTAAAGATT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATCAAGAAAAATTTTCAAAAAATTAATTAATTT 780
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAAGATCTATTAAAGATT 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
```

```
RESULT 12
AR527384
LOCUS AR527384
DEFINITION Sequence 8 from patent US 6723545.
ACCESSION AR527384
VERSION AR527384.1 GI:53914405
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 8 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX; Location/Qualifiers
FEATURES
    source 1..876
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
    AR609663 876 bp DNA linear PAT 15-DEC-2004
    Sequence 8 from patent US 6825019.
```

```
Pred. No.: 2,34e-188 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: Gaps: 0
US-10-734-719-14 (1-291) x AR527384 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGCTTATTTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGTATGTTATAGATGTAATCAATTTTATTTTGAAGATAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTTAAACATTTAAATCCAAATATCAAGAAATATGAGACCGCAACTAAATATGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAGAACTTTTACGATTAATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTCATATGTTGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATCAAGAAAAATTAATACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAAGATCTATTAAAGATT 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
RESULT 13
AR609663
LOCUS AR609663
DEFINITION Sequence 8 from patent US 6825019.
```

Db	661	ATAGA	CTAGCCG	CAAAATTTTAAATTC	CAAAATTTTATCAT	CAAGAAAAATTA	CTAC	720			
Qy	241	LysAsp	IleLeu	IleProSer	SerGly	AlaTyr	GlyLysPheSer	LysAsnIleAsn	260		
Db	721	AAAGAT	ATAC	TACTAC	TCTTCT	TAGT	GAGGCTTATG	GAAAAATTTTCA	780		
Qy	261	LysLys	IleLys	IleLys	GluAsn	ValTyr	TyrLys	LeuIleLysAsp	280		
Db	781	AAAAAA	TAAAAAT	TAAAGAAAAAT	TGTTTAT	TATACAA	AGTTTGATA	AAAAAGATCTAT	840		
Qy	281	ProSer	AspIle	LysHis	TyrPhe	LysGly	Lys	291			
Db	841	CCTAGT	GATATA	AAAGCAT	TATTTCAA	AGGAAAA	873				
RESULT 14											
AR689941											
LOCUS	AR689941				876 bp	DNA	linear	PAT 12-SEP-200			
DEFINITION	Sequence 8 from patent US 6905867.										
ACCESSION	AR689941										
VERSION	AR689941.1				GI:74471949						
KEYWORDS											
SOURCE	Unknown.										
ORGANISM	Unknown.										
REFERENCE	1 (bases 1 to 876)										
AUTHORS	Gilbert,M. and Wakarchuk,W.										
TITLE	Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity										
JOURNAL	Patent: US 6905867-A 8 14-JUN-2005; National Research Council of Canada; Ottawa; CAN;										
FEATURES											
source	1..876										
ORIGIN	/organism="unknown"										
	/mol_type="genomic DNA"										
Alignment Scores:											
Pred. No.:	2,34e-188							876			
Score:	1559.00							290			
Percent Similarity:	99.7%							Conservative:	0		
Best Local Similarity:	99.7%							Mismatches:	1		
Query Match:	99.7%							Indels:	0		
DB:	2							Gaps:	0		
US-10-734-719-14 (1-291) x AR689941 (1-876)											
Qy	1	MethLys	ValIle	IleAla	GlyAsn	GlyPro	SerLeu	LysGlu	IleAspTyrSer	20	
Db	1	ATGAAA	AAAGTTAT	TATTTGCT	TGGAATGG	CAAGTTTAA	AGAAATTTGATTAT	TTCAAG	60		
Qy	21	LeuPro	AsnAsp	PheAsp	ValPhe	ArgCys	AsnGln	PheTyr	PheGlu	AspLysTyr	40
Db	61	CTAC	CAATGATTTTGAT	TGTTAG	TATGTAATCA	AAATTTTATTTTGA	GATAAATAC	TAT	120		
Qy	41	LeuGly	LysLys	CysLys	ThrVal	PheTyr	ThrPro	AsnPhe	PheGlu	GlnTyrTyr	60
Db	121	CTTGGT	AAAAA	AAATCA	AGAGCGTGT	TTTAC	ACCCCTAATTTCTCT	TTGAGCA	AACTAC	180	
Qy	61	ThrLeu	LysHis	LeuIle	GlnAsn	GlnGlu	TyrGlu	ThrGlu	LeuIle	MetCys	80
Db	181	ACTTTAA	ACATTTTAA	TCCAAAA	TCAAGATAT	GAGCCG	AACTAATTTAT	TGTGTTCT	TAAT	240	
Qy	81	TyrAsn	GlnAla	HisLeu	GluAsn	GluAsn	PheVal	LysThr	PheTyr	AspTyr	100
Db	241	TACA	CCAAGCTCAT	CTAG	AAAAATG	AAAAATTTTGT	TAAAA	AACTTTTAC	GATTTAT	TTTCT	300
Qy	101	AspAla	HisLeu	GlyTyr	AspPhe	LysGln	LeuLys	GluPhe	AsnAla	TyrPhe	120
Db	301	GATC	CTCATTTGG	GATATGATTTTTTTT	TAAAC	ACACTTAA	AGAAATTTTAA	GCTTTAT	TTTAA	360	
Qy	121	PheHis	GluIle	terPhe	AsnGln	ArgIle	thrSer	GlyVal	TyrMet	CysAla	140

```
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATCGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATTCCTTTTAAATAGCCCTGATTTTAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLysLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCACTCTCTAGTAGGCTTATGGAAATTTTCAAAAAATATATATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATTTAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 15
LOCUS AR691837 876 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 8 from patent US 6911337.
ACCESSION AR691837
VERSION AR691837.1 GI:74479548
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert M. and Wakarchuk, W.W.
TITLE Nucleic acids encoding beta-1,4-GalNac transferase
JOURNAL Patent: US 6911337-A 8 28-JUN-2005;
National Research Council of Canada, Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,34e-188 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-14 (1-291) x AR691837 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATCAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
```

```
Db 61 CTACCAAAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAAATACTAT 120
Qy 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCTAAATTTCTTCTTTGACCAATACTACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATAATATATGTTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTAATTTTCCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGCTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATCGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCACTCTCTAGTAGGCTTATGGAAATTTTCAAAAAATATATATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATTTAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
```

Search completed: August 21, 2006, 17:38:33  
Job time : 4781.5 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 14:18:17 ; Search time 519 Seconds

(without alignments)  
5863.943 Million cell updates/sec

Title: US-10-734-719-14

Perfect score: 1564

Sequence: 1 MKKVIIAGNGPSLKEIDYSR.....KLIKDLLRLPSDKHYFKGK 291

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abss/ABSSWEB spool/US10734719/runat 21082006 141806 12305/app\_query.fasta 1  
-DB=N Geneseq -QMT=fastap -SUFFIX=P2N.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs05p  
-USER=US10734719 @CGN 1 1 1423 @runat 21082006 141806 12305 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8:\*

1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004as:.\*  
13: geneseqn2004bs:.\*  
14: geneseqn2005s:.\*  
15: geneseqn2005bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	100.0	873	8	ABT13671 C. jejuni
2	1559	99.7	876	3	AAAS3726 Campyloba
3	1559	99.7	876	8	ABT13669 C. jejuni

4	1553	99.3	873	8	ABT13670
5	1524	97.4	876	3	AAAS3724
6	1524	97.4	876	3	ABT13667
7	1521	97.3	876	3	AAAS3721
8	1521	97.3	876	8	ABT13666
9	1521	97.3	876	13	ADU77566
10	1521	97.3	11474	3	AAAS3720
11	1521	97.3	11474	8	ABT13665
12	1521	97.3	12219	13	ADU77595
13	1510	96.5	876	3	AAAS3725
14	1510	96.5	876	8	ABT13668
c	1510	96.5	876	8	ABT13668
15	744.5	47.6	191996	13	ADT05647
16	739	47.3	1293	2	AAZ25693
17	730.5	46.7	13379	13	ADT05530
18	730.5	46.7	117576	13	ADT05736
19	730.5	46.7	349980	13	ADT05648
20	721.5	46.1	110000	2	AAZ42063_03
21	692.5	44.3	4277	13	ADT05421
22	688.5	44.0	909	8	ACC71694
23	123	7.9	6128	6	ABQ67039
24	123	7.9	6129	6	ABK31236
25	123	7.9	6129	6	ABL70537
26	123	7.9	6129	6	AAAS61149
27	118	7.5	32392	6	ABL56203
28	116	7.4	4838	2	AAV70891
29	114.5	7.3	29993	10	ADB37662
30	114.5	7.3	186854	14	ADX38909
31	114	7.3	1179	3	AAAS5996
32	113.5	7.3	2375	3	AAZ58056
33	113	7.2	33521	12	ADN61337
34	113	7.2	110000	6	ABA92787_2
35	112.5	7.2	110000	2	AAT58840_4
36	112	7.2	4810	2	AAV70887
37	112	7.2	50000	6	ABL56201
38	111.5	7.1	1866	2	AAAS99653
39	111.5	7.1	2727	8	ACA23131
40	111.5	7.1	7758	6	ABL33102
41	111.5	7.1	110000	2	AAZ0248_03
42	111	7.1	5454	3	AAV70236
43	110.5	7.1	11474	3	AAAS3720
44	110.5	7.1	11474	8	ABT13665
45	110.5	7.1	50000	6	ABL55643

#### ALIGNMENTS

##### RESULT 1

ABT13671  
ID ABT13671 standard; DNA; 873 BP.

XX  
AC ABT13671;

XX  
DT 07-FEB-2003 (first entry)

XX  
DE C. jejuni bifunctional sialtransferase cstII coding sequence #6.

XX  
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.

XX  
OS Campylobacter jejuni.

XX  
PN WO200274942-A2.

XX  
PD 26-SEP-2002.

XX  
PF 22-FEB-2002; 2002WO-CA000229.

XX  
PR 21-MAR-2001; 2001US-00816028.

XX

PA (CANA ) NAT RES COUNCIL CANADA.  
 XX Gilbert M, Wakarchuk WW;  
 PI WPI; 2003-040554/03.  
 DR P-PSDB; ABU18485.  
 XX  
 PT New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Disclosure; Page 99; 107pp; English.  
 PS  
 CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumor  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention  
 XX  
 SQ Sequence 873 BP; 353 A; 117 C; 107 G; 296 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5,46e-151 Length: 873  
 Score: 1564.00 Matches: 291  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 8 Gaps: 0

US-10-734-719-14 (1-291) x ABF13671 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 DB 1 ATGAAAAAAGTTATTATGCTGCGAATGACCAAGTTAAAGAAATTGATTATCAAGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
 DB 61 CTACCAAAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
 DB 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTCTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
 DB 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATGAGCCGAATAATTATGTGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 DB 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300

QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 DB 361 TTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTACC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180

DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
 DB 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTAGAAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 DB 601 GAAAAAATTTACAAAATAAAACTATATTGCTTATGCTCTATAGTCTTTTAGCAAAATTTT 660

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
 DB 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAACAAGAAAAATAAATACTACTACT 720

QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 DB 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT 780

QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
 DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTATAGATT 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 DB 841 CCTAGTATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 2  
 AAA53726  
 ID AAA53726 standard; DNA; 876 BP.  
 XX  
 AC AAA53726;  
 XX  
 DT 15-SEP-2003 (revised)  
 DT 22-DEC-2000 (first entry)  
 XX  
 DE Campylobacter jejuni O:19 serotype CstII sialyltransferase.  
 XX  
 KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
 KW immunity; immunogen; ganglioside; ds.  
 XX  
 OS Campylobacter jejuni; O:19 serotype.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..876  
 FT /\*tag= a  
 FT /product= "CstII sialyltransferase"  
 FT  
 PN WO200046379-A1.  
 XX  
 PD 10-AUG-2000.  
 XX  
 PF 01-FEB-2000; 2000WO-CA0000086.  
 XX  
 PR 01-FEB-1999; 99US-0118213P.  
 PR 31-JAN-2000; 2000US-00495406.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 XX Gilbert M, Wakarchuk WW;  
 DR WPI; 2000-524418/47.  
 DR P-PSDB; AAY97212.  
 XX  
 PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 PT reagents and as immunogen for producing antibodies.  
 XX  
 PS Disclosure; Page 96; 120pp; English.  
 XX

CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesizing sialylated oligosaccharides such as ganglioside,  
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyze the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 15-SEP-2003 to standardise OS field)  
 XX

SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.79e-150 Length: 876  
 Score: 1559.00 Matches: 290  
 Percent Similarity: 99.7% Conservative: 0  
 Best Local Similarity: 99.7% Mismatches: 1  
 Query Match: 99.7% Indels: 0  
 DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x AAA53726 (1-876)

Qy 1 MetLysLysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 Db 1 ATGAAAAAAGTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTATTCACAGG 60  
 Qy 21 LeuProAsnAspPheAspValPheAArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
 Db 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTGAAGATAATACTAT 120  
 Qy 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
 Db 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTTGAGCAATACTAC 180  
 Qy 61 ThrLeuLysHisLeuLleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
 Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240  
 Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 Db 241 TACAACCAAGCTCATCTAGAAAATGAANAATTTGTAAAAAATTTTACGATTTATTTCT 300  
 Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
 Qy 121 PheHisGluIleTyrPheAsnGlnAArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 Db 361 TTTTACGAAATTTATTTCAATCAAGAATAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
 Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480  
 Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
 Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
 Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

Db 541 GATCGCTCGCACTATATCGGCATAGTAAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
 Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 Db 601 GAAAAAACTTACAAAAATAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 560  
 Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIlelleGlnGluLysAsnAsnTyrThr 240  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720  
 Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780  
 Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu 280  
 Db 781 AAAAAAATAAAATTTAAAGAAAAATGTTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 840  
 Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 Db 841 CCTAGTGATATAAGCAATTATTTCAAGGAAAA 873

## RESULT 3

ABT13669  
 ID ABT13669 standard; DNA; 876 BP.

AC ABT13669;

XX 07-FEB-2003 (first entry)

XX C. jejuni bifunctional sialtransferase cstII coding sequence #4.

XX Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNac transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.

XX Campylobacter jejuni.

XX WO200274942-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-CA000229.

XX 21-MAR-2001; 2001US-00816028.

XX (CANA ) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2003-040554/03.

XX P-PSDB; ABJ18482.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.

XX Disclosure; Page 98; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNac (N-  
 CC Acetyl-galactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune

CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention

XX  
SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.79e-150 Length: 876  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
Gaps: 0

US-10-734-719-14 (1-291) x ABT13670 (1-876)

QY 1 MetLysLeuValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTCGCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
QY 41 LeuGlyLysLysCysLysThrValPheTyrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAATAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTATCCAAATCAAGAAATAGACCGCACTAATTAATGCTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTTATTTTCC 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCGGCATATTCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GlyLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAAATATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTTAAATTTCAAAATTTTATCATCAAGAAAAATACTACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAGATATACTCATACCTCTTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

## RESULT 4

ABT13670  
ID ABT13670 standard; DNA; 873 BP.

XX AC ABT13670;

XX DT 07-FEB-2003 (first entry)

XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #5.

XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase; sialic acid synthase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthetase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.

XX PN WO200274942-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-CA000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX PA (CANA ) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX P-PSDB; 2003-040554/03.

XX DR P-PSDB; ABJ18484.

XX PT New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.

XX PS Disclosure; Page 98-99; 107pp; English.

XX CC The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
CC Acetylgalactosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention

XX SQ Sequence 873 BP; 350 A; 118 C; 110 G; 295 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.37e-150 Length: 873  
Score: 1553.00 Matches: 289  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 2  
Query Match: 99.3% Indels: 0  
Gaps: 0

US-10-734-719-14 (1-291) x ABT13670 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
|||||

Db 1 ATGAAAAAGTTATTATTGCTGGAATGGCAAGTTTAAAGAAATGATTATTCACAGG 60  
Qy 21 LeuProAsnAspPheAspValPheAsgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTGGTTCTTCTTTGACATACTAC 180  
Qy 61 ThrLeuLysHisLeuLeGlnAsnGlnGluTyrGluThrGluLeuLeMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAAATATGTGTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAACCACTTTTACGATTAATTTCCCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAGAAATTAATGCTTATTTTAAA 360  
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAACTAACAGTCTTTTAGCAAAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleLeuGlnLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAGATATACTCATACCTTCTAGTGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAAATTTAAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 5  
AAAS3724  
ID AAAS3724 standard; DNA; 876 BP.  
XX  
AC AAAS3724;  
XX  
DT 15-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.  
XX  
Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;

KW immunity; immunogen; ganglioside; ds.  
XX  
OS Campylobacter jejuni; O:10 serotype.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..876  
FT /\*tag= a  
FT /product= "alpha-2,3-sialyltransferase"  
XX  
PN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CA000086.  
XX  
PR 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
PA (CANA ) NAT RES COUNCIL CANADA.  
XX  
PI Gilbert M, Wakarchuk WW;  
XX  
DR WPI; 2000-524418/47.  
DR P-PSDB; RAY97210.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
PS Claim 6; Page 92; 120pp; English.  
XX  
CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
XX  
SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.05e-147 Length: 876  
Score: 1524.00 Matches: 284  
Percent Similarity: 97.9% Conservative: 1  
Best Local Similarity: 97.6% Mismatches: 6  
Query Match: 97.4% Indels: 0  
DB: Gaps: 0

US-10-734-719-14 (1-291) x AAAS3724 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGCAAGTTTAAAGAAATGATTATTCACAGG 60  
Qy 21 LeuProAsnAspPheAspValPheAsgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40



```
Db      61  CTACCAAAATGATTTTTCATGTAATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120
QY      41  LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      121  CTTGGTAAAAAATTCAAAGCAGTAATTTTCAATCTCGTCTTTTGTGTTTGAACAATACTAC 180
QY      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db      181  ACTTTAAACATTTAATCAAAATCAAGAAATATGAGACCGAACTAATATGTTCTTAAT 240
QY      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTAATTTTCCT 300
QY      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360
QY      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTTCAGAAATTTATCTCAATCAAGAAATTAACCTCAGGAGTCTATATGTCGAGTAGCT 420
QY      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATGATTTTTATCAAAATGGTCA 480
QY      161  SerTyrAlaPheAspThrLysGlnGlnAsnLeuLeuLysLeuAlaProaspPheLysAsn 180
Db      481  TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTGGCTCTGCTGATTTTAAAAAT 540
QY      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTAACAAATTAACATATATGCTATGTCCTTAACAGTCTTTTGAACAATTTT 660
QY      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGlnLysAsnAsnTyrThr 240
Db      661  ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAAATAACTACACT 720
QY      241  LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721  AAAGATATATCTCATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
QY      261  LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781  AAAAAATTAATAATTAAGAAATATATTTATACAGTTGTATATAAAGATCTATTAAAGATTA 840
QY      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841  CCTAGTATATAAGCATTATTTTCAAGGAAAA 873

RESULT 6
ABT13667
ID      ABT13667 standard; DNA; 876 BP.
XX
AC      ABT13667;
XX
DE      07-FEB-2003 (first entry)
XX
C. jejuni bifunctional sialtransferase cstII coding sequence #2.
KW      Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW      GalNAc transferase; N-Acetylglucosamine transferase;
KW      galactosyltransferase; sialyltransferase; sialic acid synthase;
KW      cytidine 5'-monophosphate sialic acid synthetase;
KW      CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW      ganglioside mimetics; inflammation; tumour metastasis.
XX
OS      Campylobacter jejuni.
```

```
PN      WO200274942 -A2.
XX
PD      26-SEP-2002.
XX
PF      22-FEB-2002; 2002WO-CA000229.
XX
PR      21-MAR-2001; 2001US-00816028.
XX
PA      (CANA ) NAT RES COUNCIL CANADA.
XX
PI      Gilbert M, Wakarchuk WW;
XX
WI      WPI: 2003-040554/03.
DR      P-PSDB; ABJ18480.
XX
PT      New glycosyltransferases from Campylobacter, useful for synthesizing
PT      gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT      mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS      Claim 8; Page 96-97; 107pp; English.
XX
CC      The invention comprises the amino acid and coding sequences of
CC      Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC      may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC      Acetylglucosamine) transferase; galactosyltransferase;
CC      sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC      sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC      sequences of the invention are useful for ganglioside synthesis, studying
CC      ganglioside mimetics, and for designing oligonucleotides to inhibit
CC      expression of Campylobacter enzymes involved in the biosynthesis of
CC      ganglioside mimetics that can mask the pathogen's from the host's immune
CC      system. The C. jejuni oligosaccharides of the invention may be used as
CC      diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC      metastasis). The present DNA sequence represents a Campylobacter jejuni
CC      gene of the invention
XX
SQ      Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      7,05e-147      Length:      876
Score:      1524.00      Matches:      284
Percent Similarity:      97.9%      Conservative:      1
Best Local Similarity:      97.6%      Mismatches:      6
Query Match:      97.4%      Indels:      0
DB:      Gaps:      0

US-10-734-719-14 (1-291) x ABT13667 (1-876)
QY      1  MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyr-SerArg 20
Db      1  ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
QY      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAATGATTTTGTATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120
QY      41  LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      121  CTTGGTAAAAAATTCAAAGCAGTAATTTTACAATCTCGTCTTTTGTGAAACAATACTAC 180
QY      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db      181  ACTTTAAACATTTAATCAAAATCAAGAAATATGAGACCGAACTAATATGTTCTTAAT 240
QY      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTAATTTTCCT 300
QY      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
```

```
Db 361 TTTTCCGAAATTTATCTCAATCAAGAAATTAACCTCAGGACTATATGTGTCAGTAGCT 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTCGAATTGATTTTATCAAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTGGCTCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGACATAGTAAATAACAGATATAAAAGCTTTAGAATTCTA 600
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTACACAGTCTTTAGCAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTATGGAATAATTTCAAAAAATATTAATTT 780
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATATTTATTAACAAGTTGATAAAAGATCTTATTAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAAGGAAAA 873
```

## RESULT 7

AAAS3721  
ID AAAS3721 standard; DNA; 876 BP.

XX AC AAAS3721;

DT 15-SEP-2003 (revised)

DT 22-DEC-2000 (first entry)

XX Campylobacter jejuni OH4384 CstII sialyltransferase coding sequence.

XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
XX acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
XX Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
XX sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
XX immunity; immunogen; ganglioside; ds.

XX Campylobacter jejuni; OH4384.

XX Key Location/Qualifiers

XX CDS 1..876

XX /\*tag= a

XX /product= "CstII sialyltransferase"

XX W0200046379-Al.

XX 10-AUG-2000.

XX 01-FEB-2000; 2000WO-CA000086.

XX 01-FEB-1999; 99US-0118213P.

XX 31-JAN-2000; 2000US-00495406.

XX (CANADA ) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2000-524418/47.

DR P-PSDB; AAY97204.

XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
XX reagents and as immunogen for producing antibodies.

PS Claim 6; Page 90; 120pp; English.

XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyze the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 15-SEP-2003 to standardise OS field)

XX SQ Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.43e-146 Length: 876  
Score: 1521.00 Matches: 282  
Percent Similarity: 98.3% Conservative: 4  
Best Local Similarity: 96.9% Mismatches: 5  
Query Match: 97.3% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x AAAS3721 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAAGTTATTATTGCTGGAATCGCAAGATTTAAAGAAATTTGATTATTCAAGA 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTGTATGATTTAGATGTTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAATGCAAGGCAGTATTTTACAACTCTATTCCTTTTGAACAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTAAATCCAAAATCAAGATATGAGACCGAACTAATATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTCAAAACACTTAAAGATTTTAAATGCTTATTATAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTCACGAAATTTTATTTCAATCAAAAGAAATTAACCTCAGGGGTTTATATGTGTCAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

```

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTAAAAATGGCTCCTAAATTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTACAAAATAAACAATATATGCTTATGTCCTAACAGTCTTTTACGAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGGCCCAATTTAAATTTTAAATTTTATCATCAAGAAAAATAACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTCATACCTTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAAGAAATATTTTATCAAGTTGATTAAGAAATCTATTAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTTCAAGGAAAA 873

RESULT 8
ABT13666
ID ABT13666 standard; DNA; 876 BP.
XX
AC ABT13666;
XX
DT 07-FEB-2003 (first entry)
XX
DE C. jejuni bifunctional sialtransferase cstII coding sequence #1.
XX
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
OS Campylobacter jejuni.
XX
PN WO200274942-A2.
XX
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-CA000229.
XX
PR 21-MAR-2001; 2001US-00816028.
XX
PA (CANADA) NAT RES COUNCIL CANADA.
XX
PI Gilbert M, Wakarchuk WW;
XX
DR WPI; 2003-040554/03.
XX
DR P-PSDB; ABJ18479.
XX
PT New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS Claim 8; Page 96; 107pp; English.
XX
CC The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention

```

```

CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present DNA sequence represents a Campylobacter jejuni
CC gene of the invention
XX
SQ Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1-43e-146 Length: 876
Score: 1521.00 Matches: 282
Percent Similarity: 98.3% Conservative: 4
Best Local Similarity: 96.9% Mismatches: 5
Query Match: 97.3% Indels: 0
DB: 8 Gaps: 0

```

US-10-734-719-14 (1-291) x ABT13666 (1-876)

```

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGGAATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGGCAGTATTTTACAATCTTCTTTTGTGAAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTATGTGTTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAAACAGCTCATCTAGAAATGAAATTTTGTAAGAACTTTTACGATTATTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTCAAAACAACCTTAAAGATTTTAATGCTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTTCAGGGGTTTATATGTGTGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTAAAAATGGCTCCTAAATTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATAACAGATATAAAGCTTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTACAAAATAAACAATATATGCTTATGTCCTAACAGTCTTTTACGAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGGCCCAATTTAAATTTTAAATTTTATCATCAAGAAAAATAACTACACT 720

```

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Dd 721 AAAGATATCTCATACCTCTAGTGGCTTATCGAAATTTTCAAAAAATATTAATTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Dd 781 AAAAAAATAAAAAATAAGAAAAATATTTATACAAAGTTGATAAAAGATCTATTAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Dd 841 CCTAGTGATATAAGCAATTTATTTCAAGAGAAA 873

## RESULT 9

ADU77566  
ID ADU77566 standard; DNA; 876 BP.  
XX AC ADU77566;  
DT 10-FEB-2005 (first entry)  
XX Sialytransferase (cstII).  
XX gene amplification; campylobacter infection; Sialytransferase; cstII; ds.  
XX Campylobacter jejuni.  
XX W02004101817-A1.  
XX 25-NOV-2004.  
XX 19-MAY-2004; 2004WO-AU000654.  
XX 19-MAY-2003; 2003AU-00902445.  
XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
XX Shi F, Fry BN, Coloe PJ;  
DR WPI; 2004-834013/82.  
XX  
PT Identifying Campylobacter in a sample to comprises comparing the profile  
PT of an amplification product to a comparator Campylobacter species after  
PT performing nucleic acid amplification of the wla gene cluster.  
XX Claim 6; SEQ ID NO 13; 75pp; English.

XX The invention describes a method of identifying Campylobacter in a sample

CC comprising comparing the profile of an amplification product to a  
CC comparator Campylobacter species after performing nucleic acid  
CC amplification. Identifying Campylobacter in a sample comprises: (a)  
CC performing nucleic acid amplification by contacting the sample with a  
CC pair of nucleic acid primers, where at least one nucleic acid primer is  
CC derived from the wla gene cluster of Campylobacter, for a time and under  
CC conditions for generation of an amplification product comprising the wla  
CC gene cluster of Campylobacter or its portion; (b) preparing a profile of  
CC the amplification product; and (c) comparing the profile to a profile of  
CC a comparator Campylobacter species, strain or type, and thus, determining  
CC the identity of Campylobacter in the sample. Also described are: (a) a  
CC pair of nucleic acid primers, where at least one nucleic acid primer is  
CC derived from the wla gene cluster of Campylobacter; or (b) a pair of  
CC nucleic acid primers for amplification of the wla gene cluster of  
CC Campylobacter or its portion; or (c) a probe capable of hybridizing to  
CC the wla gene cluster of Campylobacter. The method and kit are useful for  
CC identifying and detecting the presence of Campylobacter in a sample, or  
CC for diagnosing infections associated with Campylobacter. This sequence  
CC represents Sialytransferase, encoded by a gene in the Campylobacter  
CC jejuni wla cluster used in the inventive method of detecting  
CC Campylobacter jejuni infection.

SQ Sequence 876 BP; 356 A; 115 C; 104 G; 301 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.43e-146 Length: 876

Score: 1521.00 Matches: 283  
Percent Similarity: 97.9% Conservative: 2  
Best Local Similarity: 97.3% Mismatches: 6  
Query Match: 97.3% Indels: 0  
Dd: 13 Gaps: 0

US-10-734-719-14 (1-291) x ADU77566 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Dd 1 ATGAAAAAGTTATTTCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Dd 61 CTACCAATGATTTTGTATTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Dd 121 CTTGGTAAAAAATTCAAAGCGAGTATTTTACAAATCCTGGTCTTTTGTGAACAAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Dd 181 ACTTTAAACAATTTAAATCCAAAAATCAAGATATGAGACCGAATTAATATGTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Dd 241 TACAACCCAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Dd 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Dd 361 TTTCAACGAAATTTATCTCAATCAAGAAATTAACCTCAGAGGTATATATGTGCGATAGCT 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Dd 421 ATAGCCCTAGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Dd 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTGGCTCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Dd 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTGAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Dd 601 GAAAAAACTTACAAAAATAAAACTATATTTGCTTATGCTCTTAACAGTCTTTTACCAAAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Dd 661 ATAGAACTAGCGCCCAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACTACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Dd 721 AAAGATATCTCATACCTTCTAGTGAGCTTATGAGAAAAATTTTCAAAAAATTTAATTAATTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Dd 781 AAAAAAATAAAAAATAAGAAAAATTTATTTTACAAAGTTGATAAAAGATCTATTAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Dd 841 CCTAGTGATATAAGCAATTTATTTCAAGAGAAA 873

## RESULT 10

AAA53720  
ID AAA53720 standard; DNA; 11474 BP.  
XX  
AC AAA53720;

```

XX 15-SEP-2003 (revised)
DT 22-DEC-2000 (first entry)
XX
XX LPS core biosynthesis locus.
DE
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside; ds.
XX
XX Campylobacter jejuni; OH4384.
XX
XX Key Location/Qualifiers
FH 350..1237
CDS /*tag= a
/*product= "acyltransferase"
/*note= "Open reading frame 2b"
FT 1234..2490
CDS /*tag= b
/*product= "glycosyl transferase"
/*note= "Open reading frame 3a"
FT 2786..3955
CDS /*tag= c
/*product= "Glycosyl transferase"
/*note= "Open reading frame 4b"
FT 4025..5068
CDS /*tag= d
/*product= "Beta 1,4-GalNAC transferase"
/*note= "Open reading frame 5a"
FT 5054..5959
CDS /*tag= e
/*product= "Beta 1,2- galactosyltransferase"
/*note= "Open reading frame 6a"
FT 6048..6923
CDS /*tag= f
/*product= "CstII sialyltransferase"
/*note= "Open reading frame 7a"
FT 6924..7964
CDS /*tag= g
/*product= "Sialic acid synthase"
/*note= "Open reading frame 8a"
FT 8020..9079
CDS /*tag= h
/*product= "Sialic acid biosynthetic enzyme"
/*note= "Open reading frame 9a"
FT 9076..9741
CDS /*tag= i
/*product= "CMP-sialic acid synthetase"
/*note= "Open reading frame 10a"
FT complement(10554..11366)
CDS /*tag= j
/*product= "Lipooligosaccharide biosynthetic enzyme"
/*note= "Open reading frame 12a"
FT
XX
XX WO2000046379-A1.
XX
XX 10-AUG-2000.
XX
XX 01-FEB-2000; 2000WO-CA0000086.
XX
XX 01-FEB-1999; 99US-0118213P.
PR 31-JAN-2000; 2000US-00495406.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
PI
XX WPI: 2000-524418/47.
DR P-PSDB; AAY97200, AAY97201, AAY97202, AAY97203, AAY97204, AAY97205,
DR AAY97206.
XX

```

```

PT Novel glycosyltransferase polypeptides and polynucleotides useful for
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
PT reagents and as immunogen for producing antibodies.
XX
XX PS Claim 1; Page 86-90; 120pp; English.
XX
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is
CC useful for synthesising sialylated oligosaccharide such as ganglioside,
CC lysoanglioside or their mimics. Glycosyltransferases are useful for
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
CC other oligosaccharides that have biological activity. The enzymes and
CC nucleic acids that encode them are useful for studies of the pathogenesis
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
CC jejuni and the nucleic acids are used as probes to study expression of
CC genes involved in ganglioside mimetic synthesis. Antibodies raised
CC against the glycosyltransferases are also useful for analyzing the
CC expression patterns of these genes involved in pathogenesis. The nucleic
CC acids are also useful for designing antisense oligonucleotides for
CC inhibiting expression of the Campylobacter enzymes that are involved in
CC the biosynthesis of ganglioside mimics that can mask the pathogens from
CC the host's immune system. The oligosaccharides are useful as diagnosing
CC reagents or as therapeutics and as immunogens for producing antibodies.
CC Bacterial glycosyltransferase can be used to catalyse the formation of
CC oligosaccharides that are identical to the corresponding mammalian
CC structures and are easier and less expensive to produce in large
CC quantity compared to the mammalian glycosyltransferase. The bacterial
CC origin of the enzymes facilitates expression of large quantities of the
CC enzymes using relatively inexpensive prokaryotic expression systems.
CC (Updated on 15-SEP-2003 to standardise OS field)
XX
XX SQ Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.85e-145 Length: 11474
Score: 1521.00 Matches: 282
Percent Similarity: 98.3% Conservative: 4
Best Local Similarity: 96.9% Mismatches: 5
Query Match: 97.3% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x AAAS3720 (1-11474)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGGAATAAATACTAT 6167
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6168 CTTGGTAAAAAATGCAAGGCAGTATTTTACAATCCTATTCTTTTGTGAAACAATAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTTAAACATTTTAATCCAAAATCAAGAAATATGAGACCGAATTAATATGTGTTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATATGATTTTTCACAACTTAAGATTTTAAATGCTTATTATTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCCAGAAATTTTATTTCATCAATCAAAAGAAATACCTCAGGGGCTTATATGTGCTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTGATTTTATCAAAATGGGTCA 6527

```

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
|||||  
Db 6528 TCITATGCTTTTGATACATAACAAATACTTTTAAATGGCTCCTCTAATTTTAAAAAT 6587  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200  
|||||  
Db 6588 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTA 6647  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
|||||  
Db 6648 GAAAAAATCTACAAAAATAAACATATATGCTTATGCTTAAACAGCTCTTTAGCAAAATTTT 6707  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnThrTyr 240  
|||||  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT 6767  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
|||||  
Db 6768 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827  
Qy 261 LysLysIleLysLysLeuGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
|||||  
Db 6828 AAAAAAATAAAAAATTAAGAAAAATATTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 6887  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
|||||  
Db 6888 CCTAGTATATAAGCAATATTTTCAAAAGGAAAA 6920  
RESULT 11  
ABTI13665  
ID ABTI13665 standard; DNA; 11474 BP.  
XX  
AC ABTI13665;  
XX  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni genomic lipooligosaccharide biosynthesis locus.  
XX  
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GAINAC transferase; N-Acetylglactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.  
XX  
XX  
XX WO200274942-A2.  
XX  
PD 26-SEP-2002.  
XX  
XX  
PF 22-FEB-2002; 2002WO-CA000229.  
XX  
XX  
PR 21-MAR-2001; 2001US-00816028.  
XX  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
PI  
XX  
XX WPI; 2003-040554/03.  
DR  
XX  
XX New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
XX  
PS Claim 1; Page 89-95; 107pp; English.  
CC  
XX The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GAINAC (N-  
CC Acetylglactosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)

CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention  
XX  
SQ Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.65e-145 Length: 11474  
Score: 1521.00 Matches: 282  
Percent Similarity: 98.3% Conservative: 4  
Best Local Similarity: 96.9% Mismatches: 5  
Query Match: 97.3% Indels: 0  
DB: 8 Gaps: 0  
US-10-734-719-14 (1-291) x ABTI13665 (1-11474)  
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 6048 ATGAAAAAGTTATTATTGCTGGAAATCGACCAAGCTTTAAAAAGAAATGATTATTCAAGA 6107  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 6108 CTACCAATGATTTTGTATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 6168 CTTCGTAAAAATGCAAGGCAGTATTTTACAATCTATTCTTTTGTGAACAATACTAC 6227  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 6228 ACTTTAAACATTTAATCCAAAAATCAAGATATGAGACCGAACTAATATATGTGTTCTAAT 6287  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTCCT 6347  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 6348 GATGCTCATTTGGGATATGATTTTTCACACAACTTAAAGATTTTAAAGCTTATTTTAA 6407  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 6408 TTTCAAGAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 6467  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTTATCAAAATGGGTCA 6527  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 6528 TCTTATGCTTTTGATACTATAACAAAAATCTTTTAAAAATTCGCTCCTTAATTTAAAAAT 6587  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 6588 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTA 6647  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 6648 GAAAAAATCTACAAAAATAAACTATATGCTTATGCTTAAACAGCTCTTTTACCAAAATTTT 6707  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnThrTyr 240  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACTACT 6767  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 6768 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827

QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
 Db 6828 AAAAAAATAAAAAATAAGAAAAATATTTATACAGTTGATAAAAGATCTATTAAAGATTA 6887  
 QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 Db 6888 CCTAGTGATATAAAGCATATTTCAAAGGAAAA 6920

RESULT 12  
 ADU77595  
 ID ADU77595 standard; DNA; 12219 BP.  
 XX  
 AC ADU77595;  
 XX  
 XX 10-FEB-2005 (first entry)  
 XX  
 XX Campylobacter jejuni strain ATCC 43469 wla cluster.  
 XX  
 XX gene amplification; campylobacter infection; wla; ds.  
 XX  
 XX Campylobacter jejuni.  
 XX  
 XX WO2004101817-A1.  
 XX  
 XX 25-NOV-2004.  
 XX  
 XX 19-MAY-2004; 2004WO-AU000654.  
 XX  
 XX 19-MAY-2003; 2003AU-00902445.  
 XX  
 XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
 PA  
 XX Shi F, Fry BN, Coloe PJ;  
 XX WPI; 2004-834013/82.  
 DR  
 XX  
 PT Identifying Campylobacter in a sample to comprises comparing the profile  
 PT of an amplification product to a comparator Campylobacter species after  
 PT performing nucleic acid amplification of the wla gene cluster.  
 XX  
 XX Disclosure; Fig 6; 75pp; English.  
 PS  
 XX

The invention describes a method of identifying Campylobacter in a sample comprising comparing the profile of an amplification product to a comparator Campylobacter species after performing nucleic acid amplification. Identifying Campylobacter in a sample comprises: (a) performing nucleic acid amplification by contacting the sample with a pair of nucleic acid primers, where at least one nucleic acid primer is derived from the wla gene cluster of Campylobacter, for a time and under conditions for generation of an amplification product comprising the wla gene cluster of Campylobacter or its portion; (b) preparing a profile of the amplification product; and (c) comparing the profile to a profile of a comparator Campylobacter species, strain or type, and thus, determining the identity of Campylobacter in the sample. Also described are: (a) a pair of nucleic acid primers, where at least one nucleic acid primer is derived from the wla gene cluster of Campylobacter; or (b) a pair of nucleic acid primers for amplification of the wla gene cluster of Campylobacter or its portion; or (c) a probe capable of hybridizing to the wla gene cluster of Campylobacter. The method and kit are useful for identifying and detecting the presence of Campylobacter in a sample, or for diagnosing infections associated with Campylobacter. This sequence represents a Campylobacter jejuni strain ATCC 43469 wla cluster.

SQ Sequence 12219 BP; 4778 A; 1598 C; 1745 G; 4098 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	3.95e-145	Length:	12219
Score:	1521.00	Matches:	283
Percent Similarity:	97.9%	Conservative:	2
Best Local Similarity:	97.3%	Mismatches:	6
Query Match:	97.3%	Indels:	0
DB:	13	Gaps:	0

US-10-734-719-14 (1-291) x ADU77595 (1-12219)  
 QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 Db 5995 ATGAAAAAGTTATTATTTCTGGAATGGACCAAGTTTAAAGAAAAATGATTATTCAAG 6054  
 QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
 Db 6055 CTACCAAAATGATTTTGTATGTAATGATGCAATCAATTTTATTTTGAAGATAAATACTAT 6114  
 QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
 Db 6115 CTGTGTAATAAAATTCAAAGCAGATTTTACAATCCTGGTCTTTTGTGAACATACTAC 6174  
 QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
 Db 6175 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCCGAACCTAATATATGTGTCTAAT 6234  
 QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 Db 6235 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATATTATTCCT 6294  
 QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 Db 6295 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTAAA 6354  
 QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 Db 6355 TTTTCAGCAATTTATCTCAATCAAAGAAATTAACCTCAGGAGCTATATGTGTGAGTAGCT 6414  
 QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 Db 6415 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAATTCATTTTATCAAAATGGGTCA 6474  
 QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
 Db 6475 TCTTATGCTTTTGATACCAACAAGAAAATCTTTAAAACCTGGCTCCTGATTTAAAAAT 6534  
 QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
 Db 6535 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 6594  
 QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 Db 6595 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTTAACTAACAGTCTTTTAGCAAAATTT 6654  
 QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
 Db 6655 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACTACT 6714  
 QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 Db 6715 AAAGATATACTCATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAAATTTT 6774  
 QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
 Db 6775 AAAAAAATAAAAAATAAGAAAAATATTATTACAGTTGATAAAAGATCTATTAAAGATTA 6834  
 QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 Db 6835 CCTAGTGATATAAAGCATATTTCAAAGGAAAA 6867

RESULT 13  
 AAA53725  
 ID AAA53725 standard; DNA; 876 BP.  
 XX  
 AC AAA53725;  
 XX  
 XX 15-SBP-2003 (revised)  
 DT 22-DEC-2000 (first entry)  
 XX  
 DE Campylobacter jejuni O:41 serotype alpha-2,3-sialyltransferase.

XX Biosynthetic locus; glycosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside; ds.  
XX  
OS Campylobacter jejuni; O:41 serotype.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 1..876  
FT /\*tag= a  
FT /product= "alpha-2,3-sialyltransferase"  
XX  
XX WO2000046379-A1.  
XX  
XX 10-AUG-2000.  
XX  
XX 01-FEB-2000; 2000WO-CA000086.  
XX  
XX 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX  
XX WPI; 2000-524418/47.  
DR P-PSDB; AAY97211.  
XX  
XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
XX Claim 6; Page 94; 120pp; English.  
XX  
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 15-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1..93e-145 Length: 876  
Score: 1510.00 Matches: 278  
Percent Similarity: 98.3% Conservative: 8  
Best Local Similarity: 95.5% Mismatches: 5  
Query Match: 96.5% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x AAAS3725 (1-876)  
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyr-SerArg 20

1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAGATTTTAAAGAAATTTGATTATTCAAGA 60  
21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
61 CTACCAATGATTTTGATGATTTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120  
41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
121 CTGGTAAAAATGCAAAAGCAGTATTTTACAATCCTAGTCTTTTTTTTGAACAATACTAC 180  
61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
181 ACTTTAAACATTTTAAATCCAAAAATCAAGATATGAGACCGAATAATCATGTGTTCTAAT 240  
81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
241 TTTAACCAAGCTCATCTAGAAAAATCAAAATTTTGTAAAACTTTTACCATTATTTCCT 300  
101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
301 GATGCTCATTTGGGATATGATTTTCAACCAACTTAAAGAAATTCATGCTTATTATTAA 360  
121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
361 TTTCAAGAAATTTTATTTCAATCAAGAATTACCTCAGGGTCTATATGTGCACAGTAGCC 420  
141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGATCA 480  
161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
481 TCTTATGCTTTTGATACCAAAACAAAAATCTTTTAAAAATTGGCTCCTTAATTTTAAAA 540  
181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
541 GATTAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600  
201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
601 GAAAAAACTTACGAAATAAAGCTATATTGTTTATGCTCTAACAGTCTTTTAGCAATTTT 660  
221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
661 ATAGAACCTAGCGCCCAAAATTTAAATTTCAAAATTTTATCACAAGAAAAATACTATACT 720  
241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
721 AAAGATATCTCATCTTCTTAGGAGCTTATGGAAAAATTTACAAAAATATTATTTT 780  
261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
781 AAAAAATAAAAAATTAAGAAAAATATTATTATCAAGTTTGATAAAAGATCTATTAGATT 840  
281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
RESULT 14  
ABT13668 standard; DNA; 876 BP.  
XX  
XX ABT13668;  
XX  
XX 07-FEB-2003 (first entry)  
XX  
XX C. jejuni bifunctional sialtransferase cstII coding sequence #3.  
XX  
XX Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAC transferase; N-Acetyl-galactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;



KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 PN WO200274942-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 XX 22-FEB-2002; 2002WO-CA000229.  
 PF  
 XX 21-MAR-2001; 2001US-00816028.  
 XX  
 XX (CANADA) NAT RES COUNCIL CANADA.  
 PA  
 XX Gilbert M, Wakarchuk WW;  
 PI  
 XX WPI; 2003-040554/03.  
 XX  
 XX P-PSDB; ABJ18481.  
 XX  
 XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 XX Claim 8; Page 97; 107pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 CC Acetyl-galactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention  
 XX

SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.93e-145 Length: 876  
 Score: 1510.00 Matches: 278  
 Percent Similarity: 98.3% Conservative: 8  
 Best Local Similarity: 95.5% Mismatches: 5  
 Query Match: 96.5% Indels: 0  
 DB: 8 Gaps: 0

US-10-734-719-14 (1-291) x ABT13668 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 Db 1 ATGAAAAAGTTAATTCCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGA 60  
 Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
 Db 61 CTACCAAAATGATTTTGATGATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120  
 Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
 Db 121 CTGGTAAAAAATGCAAGCAGTATTTTACAATCTAGTCTTTTTTTTGAACAATACTAC 180  
 Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluLeuIleMetCysSerAsn 80  
 Db 181 ACTTTAAACATTTAAATCCAAAATCAAGAATATGAGACCGAACTAAATCATGTGTTCTAAT 240  
 Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluAsnGluAsnGluAsnGluAsn 100  
 Db 241 TTTTAAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAAAACTTTTACGATTTATTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 Db 301 GATGCTCAITTTGGGATATGATTTTTCAAACAACACTTAAAGAAATTCATGCTTATTTAAA 360  
 Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 Db 361 TTTTCAGCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGCACAGTAGCC 420  
 Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGATCA 480  
 Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
 Db 481 TCTTATGCTTTTGATACCAACAACAAAAATCTTTTAAAAATTTGGCTCCTTAATTTAAAAAT 540  
 Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
 Db 541 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
 Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 Db 601 GAAAAAATTTACGAAATAAAGCTATATTTGTTATGTCTTACAGTCTTTTACGAAAATTTT 660  
 Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
 Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATACTACTACT 720  
 Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 Db 721 AAAGATATATCTATACCTCTAGTGAGCTTATGGAATAATTTACAAAAATATTAATTTT 780  
 Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
 Db 781 AAAAAAATAAAAAATAAAGAAAAATATTTATTTACAAGTTGATAAAAAAGATCTTATTAAGATTA 840  
 Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 Db 841 CCTAGTGATATAAAGCATTTATTTTCAAGGAAAA 873  
 RESULT 15  
 ADT05647/c  
 ID ADT05647 standard; DNA; 191996 BP.  
 XX  
 AC ADT05647;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Haemophilus influenzae (NTHI) DNA sequence - SEQ ID 683.  
 XX  
 KW middle ear bacterial infection; nasopharynx bacterial infection; ds.  
 XX  
 OS Haemophilus influenzae.  
 XX  
 PN WO2004078949-A2.  
 XX  
 PD 16-SEP-2004.  
 XX  
 XX 05-MAR-2004; 2004WO-US0007001.  
 PF  
 XX 06-MAR-2003; 2003US-0453134P.  
 PR  
 XX (CHIL-) CHILDRENS HOSPITAL INC.  
 PA  
 XX Bakletz LO, Munson RS, Dyer DW;  
 PI  
 XX WPI; 2004-662422/64.  
 DR  
 XX New polynucleotides of nontypeable strain of Haemophilus influenzae,  
 PT useful for treating or preventing NTHi bacterial infections of the middle  
 PT ear and/or nasopharynx.  
 XX  
 XX Claim 1; SEQ ID NO 683; 88pp; English.  
 PS





High quality sequence stop: 794.		Location/Qualifiers		1..907		/organism="Entamoeba histolytica"		/mol_type="genomic DNA"		/strain="HMI:IMSS"		/db_xref="taxon:5759"		/clone_lib="Entamoeba histolytica Sheared DNA"		Institute for Genomic Research (TIGR), Rockville, MD		Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G. and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999)."																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
FEATURES		source		5,16e-05		Length: 907		Score: 142.00		Matches: 72		Percent Similarity: 36.9%		Conservative: 48		Best Local Similarity: 22.2%		Mismatch: 93		Query Match: 9.1%		Indels: 112		DB: 11		Gaps: 15																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
ORIGIN		US-10-734-719-14 (1-291) x AZ685326 (1-907)		38		LysTyrTyrLeuGlyLys-----LysCysLysThrValPheTyrThrProAsnPhePhe		55		1		CAGTTTGGTCTGTTAAAGAAAGTAAATGTCACCTCTGAAACAGCTCTGTAAGCGCATTA		60		56		PheGluGlnTyrThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeu		75		61		GTGATGAGTTCATTTTAAAGAACTA-----AAAGACTTTGAAATTTGAAAT		111		76		IleMetCysSer---AsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThr		94		112		AGAAATGCTAGTACTAATATTATTTCAAAATTTTATTGAAATAAAGGAATA-----		162		95		PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu		114		163		-----GATGCT-----TTCTTTAAATGTTTAAATAA		189		115		-----PheAsnAlaTyrPheLysPheHisGluIle-----		124		190		CTTATGATGAACCTCATTTTAAAGTTACAGAAATTCCTCAAGCTGAAAGAAACATATT		249		125		-----TyrPheAsnGlnArgIleThrSerGly		133		250		TATATTTTAAAGCGCTTTTAGAACTGAAACAATTTTAAATAACCTTATTGTAGTGCA		309		134		ValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle-----		148		310		GAATATCTTGTACTATTTTAAATGGCATTTATCTCTTAAGTCTATCTCTACAAACGAACGA		369		149		-----TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerTyr-----		162		370		ATTGAATTAAGTTAATGAGCTCTTACTAAATCATGTGGTAATCTCTGAAGAAACATACAACT		429		163		---AlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp		181		430		CCACCTTATGAATAAATTTCTAAATAATTAATTAAGCATCTACTAGCTTAAAGTAA		489		182		ArgSerHisTyrIleGly-----HisSerLysAsnThrAsp		193		490		AGAAAAATAAATTAAGATTGATGCTGTGTTAACTTTTGTATACAGAAATAATGAA		549																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
FEATURES		source		107		108		109		110		111		112		113		114		115		116		117		118		119		120		121		122		123		124		125		126		127		128		129		130		131		132		133		134		135		136		137		138		139		140		141		142		143		144		145		146		147		148		149		150		151		152		153		154		155		156		157		158		159		160		161		162		163		164		165		166		167		168		169		170		171		172		173		174		175		176		177		178		179		180		181		182		183		184		185		186		187		188		189		190		191		192		193		194		195		196		197		198		199		200		201		202		203		204		205		206		207		208		209		210		211		212		213		214		215		216		217		218		219		220		221		222		223		224		225		226		227		228		229		230		231		232		233		234		235		236		237		238		239		240		241		242		243		244		245		246		247		248		249		250		251		252		253		254		255		256		257		258		259		260		261		262		263		264		265		266		267		268		269		270		271		272		273		274		275		276		277		278		279		280		281		282		283		284		285		286		287		288		289		290		291		292		293		294		295		296		297		298		299		300		301		302		303		304		305		306		307		308		309		310		311		312		313		314		315		316		317		318		319		320		321		322		323		324		325		326		327		328		329		330		331		332		333		334		335		336		337		338		339		340		341		342		343		344		345		346		347		348		349		350		351		352		353		354		355		356		357		358		359		360		361		362		363		364		365		366		367		368		369		370		371		372		373		374		375		376		377		378		379		380		381		382		383		384		385		386		387		388		389		390		391		392		393		394		395		396		397		398		399		400		401		402		403		404		405		406		407		408		409		410		411		412		413		414		415		416		417		418		419		420		421		422		423		424		425		426		427		428		429		430		431		432		433		434		435		436		437		438		439		440		441		442		443		444		445		446		447	





Query Match:	8.0%	Indels:	100
DB:	6	Gaps:	20
US-10-734-719-14 (1-291) x AK165838 (1-2789)			
Qy 60 TyrThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGlu-----LeuLeuMet 77			
Db 1465 CTGCTTGAACAACACTTCATAGAAATTTGGAGTGGGAGTCAGAATATACATTTGTTATG 1524			
Qy 78 CysSer-----AsnTyrAsnGlnAlaHis-----Leu 86			
Db 1525 TCAGCCTAAAGGTGAATAAACAATGCCAAATATTGGAATATGGGTCTATGCTCTG 1584			
Qy 87 GluAsnGlu---AsnPhValLysThrPheTyrAspTyrPhe----- 99			
Db 1585 GAGAATCAGAAGAACTTTGAGAAAGCTTTG---AAATACCTTTGTCAGGCTACCCATGTT 1641			
Qy 100 ---ProAsp-----AlaHisLeuGly-----TyrAspPheLysGln 111			
Db 1642 CAGCCAGATGACATCGGTGCCCATCATGATGTAGGAAGAACTTATAAAAACTTTAAACAG 1701			
Qy 112 LeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuLeuTyrPheAsnGlnArgLeuThr 131			
Db 1702 ACTAGAGAAGCTGAAGCGCTTTCATGCTCGGTAAATACATGATGCTTCAGATATACCT 1761			
Qy 132 SerGlyValTyrMetCysAlaValAlaAlaLeuGlyTyrLysGluLeuTyrLeuSer 151			
Db 1762 GGTAAAAAATAT-----GCAGCCAGAAATGCCCTTAATCACCTAAACGTTTATATCAAT 1815			
Qy 152 GlyLeuAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGln-----Glu 169			
Db 1816 CTGGCCAACTTATTCGAGCAATAGTCCCGCTGGAGGAAGCGGACCATGTTACCGA 1875			
Qy 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis--- 188			
Db 1876 CAGCCATCAGCATGAGCCAGACCTTCAAA-----CAGCGTTACATAGCAGGGGA 1926			
Qy 189 -----SerLysAsnThrAspIleLysAlaLeu 197			
Db 1927 GAGTTGCTTTTAAAAATGAATAGCCTCTCAAGCAAGGAAGCATATCTTAAGCACTA 1986			
Qy 198 GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCys----- 212			
Db 1987 GAG---CTGGACAGAAATAATGCAGATCTCTGGTACAACTTGGCAATGTTTATATTGAA 2043			
Qy 213 -----ProAsnSerLeuLeuAlaAsnPhe-----IleGluLeuAlaProAsn 226			
Db 2044 CTTAAAGAACCAATGAAGCTCTGAAGAACTTTAATCGAGCTTTGGAACTGAACCCCTAA 2103			
Qy 227 -----LeuAsnSerAsnPheIleLeuGlnGluLysAsnAsnTyrThrLys 241			
Db 2104 CATAGAGTACGACTGTTCAATTCCTGCTATTTTAAATGACAGAAATCA----- 2148			
Qy 242 AspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLys 261			
Db 2149 -----GGTAAGTTTCTGAAAAATGCTCTCTATTAA 2178			
Qy 262 LysIleLysIle-----LysGluAsnValTyrTyrLys 272			
Db 2179 TCAATTAATAATCTGTTAAATCCCATATAAATATTTTGTGCCAGGCTGTTTATTCGAA 2238			
Qy 273 LeuIleLysAspLeuLeuArg-----LeuProSer 282			
Db 2239 ATATATACATAATACCACTGATGAATATTTAAATTAATCTGACATAAAGTCCTTCAAGT 2298			
Qy 283 AspIleLys---HisTyrPheLysGly 290			
Db 2299 GACATGAAGAGGCATATGTGGAGGTT 2325			
RESULT 4			
DU782388			
LOCUS			
DEFINITION	ASXB4195.b2 HF500 10-06-02 uncultured marine microorganism		
	linear	GSS 27-JAN-2006	

RESULT 4					
DU782388					
LOCUS	DU782388	1013 bp	DNA	linear	GSS 27-JAN-2006
DEFINITION	ASXB4195.b2 HF500 10-06-02 uncultured marine microorganism				

```
Db 385 -----TTATCATGAATGCTCCAGGACCACATAGAAAAACATGGAAGTGTTC 435
Qy 79 SerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyr 98
Db 436 GCT-----ATTGAAGGAGGATCTCTCTTTTAGAACAAAA 471
Qy 99 PheProAspAlaHisLeuGlyTyrAspPhePheLeuGlnLeuLysGluPheAsnAlaTyr 118
Db 472 GAAGAGCTATTACATTA-----AAGAAATATCAAAAAATATATTGTT 516
Qy 119 PheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAla 138
Db 517 TTT-----TATCATCTCGGATATTTTCG--AAACTTGGAAACATTTGGCGAGAGATTGTTCT 570
Qy 139 ValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr----- 156
Db 571 TTAGCAACTCACCTTCAAGTTAAGAAAAATATTTTGTGGAATGACGATATCCCGGA 630
Qy 157 -----TGATGAACATGCTTTGAAGGAGGTCGAAAGAACACACGAGGCGGA 690
Db 631 AAGGCGCGCAATTTGTATGAACATGCTTTGAAGGAGGTCGAAAGAACACACGAGGCGGA 690
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 691 ACTTTTTCGTATGATTTGCACAGAGGCAATATGTTTATATATGGGATTAATTATGAAC 750
Qy 181 AspArgSerHisTyrIle-----GlyHisSerLysAsn----- 191
Db 751 ACCTTAGGTTCCAAGGTAAATAATATCAGAAATTTAGGAGAGGCCACGACGACCAATCAGTCG 810
Qy 192 ThrAspIleLysAlaLeu-GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLe 211
Db 811 ACTGATATTATTAGACAAATATTTCCTTTGGAAATACCATAAAATA----- 856
Qy 211 uCysProAsnSerLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPh 231
Db 857 -----TGTATATTAACTTTAGTTTATATAGAGAAATTTATGAACGATTCATT 903
Qy 231 elleIleGlnLysAsn---AsnTyrThrLysAspIleLeuIleProSerSer 248
Db 904 CATTATCGAGAAATCCGTATTAAATCACAATGGGGATATCGAAATAGCAAAATCT 958

RESULT 5
CC068914/c
LOCUS
DEFINITION
  CSU-K33r.25C12.SP6 CSU-K33r Aedes aegypti genomic clone
ACCESSION
  CC068914
VERSION
  CC068914.1 GI:29907420
SOURCE
  Aedes aegypti (yellow fever mosquito)
ORGANISM
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
  Culicidae; Culicinae; Culicini; Aedes; Stegomyia.
REFERENCE
  1 (bases 1 to 837)
  Loftus,B., Shetty,J., Severson,D., Brown,S. and Knudson,D.
  End sequencing of Aedes aegypti BACs
  Unpublished (2003)
JOURNAL
  Other_GSSs: CSU-K33r.25C12.T7
COMMENT
  Contact: Brendan Loftus
  Department of Eukaryotic Genomics
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-3543
  Fax: 301-838-0208
  Email: enta@tigr.org
  Library was provided by Susan Brown and Dennis Knudson at Colorado
  State University.
  Seq primer: SP6
  Class: BAC ends.
  Location/Qualifiers
    1..837
    source
```

```
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Rexville"
/db_xref="taxon:7159"
/clone="CSU-K33r.25C12"
/clone_lib="CSU-K33r"
/notes="Vector: pBelcBAC11; Site_1: HindIII"
```

## ORIGIN

## Alignment Scores:

Pred. No.:	0.0196	Length:	837
Score:	118.50	Matches:	58
Percent Similarity:	44.6%	Conservative:	42
Best Local Similarity:	25.9%	Mismatches:	84
Query Match:	7.6%	Indels:	42
DB:	12	Gaps:	9

US-10-734-719-14 (1-291) x CC068914 (1-837)

Qy	28	PheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysThr	47
Db	633	TTTCGAAACCAAAGAATTTCTTCGAAACCAAAGAATTTCTTCGAAATCA---AAGAA	577
Qy	48	ValPheTyrThrProAsnPhePhe-----PheGluGlnTyrTyrThrLeuLysHis	64
Db	576	TTCTTTTCGAAATTCAAAGATTTTTCAAATTCAAAGAATTCGAAATTTAAAGAAT	517
Qy	65	LeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAla	84
Db	516	TTCTTTTCGAAAT-----TTAAAGAATTTACTT-TCGAAA	485
Qy	85	HisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeu	104
Db	484	TTCAAGAATTTTCATTCGAAATTCAAAGAATTTTCATTCGAAATTCAAAGAATTTCA	425
Qy	105	GlyTyr---AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys---PheHis	122
Db	424	AAATTCAAGAATTTCTTCGAAATTCAAAGAATTTTCATTCAAATTCAAAGAATTTCA	365
Qy	123	GluIleTyr-----PheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	364	TCGAAATTCAAAGAATTTTCATTCGAAA-----	338
Qy	141	IleAlaLeuGlyTyrLysGluIleTyr-LeuSerGlyIleAspPheTyrGlnAsnGlySe	160
Db	337	-----TTCAAGAATTTTCATTCGAAATTCAAAGAATTTTCATTCGAAATTCAAA	290
Qy	160	rSerTyrAlaPheAspThrLys-----GlnGluAsnLeuLeuLysLeuAl	175
Db	289	GAATTTTCATTCGAAATTCAAAGAATTTTCATTCGAAATTCAAAGAATTTTCATTCG	230
Qy	175	aProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLy	195
Db	229	AAAGAATTTTCATTCGAAATTCAAAGAATTTTCATTCGAAATTCAAAGAATTTT-CA	171
Qy	195	sAlaLeuGluPheLeuGluLysThrTyr-LysIleLysLeuTyrCysLeuCysProAsnS	215
Db	170	ATTCATAGAATTTCTTTTCGAAATTCATTCGAAATTCAAAGAATTTTCGTTTGAATTC	111
Qy	215	erLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnG	235
Db	110	AAATTTCTTTTCGAAATTTAAAGAATTTCTTTTAAATTTA---AAGAAATTTCTTCA	54
Qy	235	lulysAsn	237
Db	53	CCAAAAAT	46

## RESULT 6

DU782007

LOCUS

DEFINITION

DU782007 1020 bp DNA linear GSS 27-JAN-2006  
ASXB4011.b2 HF500.10-06-02 uncultured marine microorganism  
HF500.10-06-02 genomic clone HF0500\_101C10, genomic survey  
sequence.



```
ACCESSION DU782007
VERSION DU782007.1 GI:85796294
KEYWORDS GSS.
SOURCE uncultured marine microorganism HF500_10-06-02
ORGANISM uncultured marine microorganism HF500_10-06-02
REFERENCE unclassified sequences; environmental samples.
AUTHORS 1 (bases 1 to 1020)
DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J.,
Frigard, N.U., Martinez, A., Sullivan, M., Edwards, R., Chisholm, S.W.
and Karl, D.M.
TITLE Comparative genomics reveals ecological trends in stratified
JOURNAL microbial communities in the ocean's interior
COMMENT Science (2006) In press
Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
and Edward DeLong
US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679
Email: PMRichardson@lbl.gov; delong@mit.edu
North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
DNA library prepared from marine picoplankton in the less than 1.6
um, greater than 0.22 um fraction. Sample Date: 10/6/2002
Coordinates: 22.45 N, 158 W Depth 500 m Temperature: 7.25 C
Salinity: 34.07 psu Oxygen: 118.0 umol/kg
Class: fosmid ends.
FEATURES             Location/Qualifiers
     source            1..1020
                     /organism="uncultured marine microorganism HF500_10-06-02"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:361149"
                     /clone="HF0500_101C10"
                     /cell_type="marine picoplankton, less than 1.8 um, greater
                     than 0.22 um fraction"
                     /clone_lib="HF500_10-06-02"
                     /notes="vector: pC1FOS; North Pacific Subtropical Gyre
                     (Hawaii) picoplankton genomic fosmid DNA library prepared
                     from marine picoplankton in the less than 1.6 um, greater
                     than 0.22 um fraction. Picoplankton collected at 500 m
                     depth on 10/6/2002. Coordinates: 22.45 N, 158 W. Sample
                     Date: 10/6/2002 Coordinates: 22.45 N, 158 W Depth 500 m
                     Temperature: 7.25 C Salinity: 34.07 psu Oxygen: 118.0
                     umol/kg"
ORIGIN
Alignment Scores:
Pred. No.:          0.199          Length:          1020
Score:              110.50         Matches:         67
Percent Similarity: 38.8%          Conservative:    47
Best Local Similarity: 22.8%       Mismatches:     116
Query Match:        7.1%           Indels:         64
DB:                  14            Gaps:           14
US-10-734-719-14 (1-291) x DU782007 (1-1020)
QY      2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      223 AAAACTATTCATGCTAGGTGGAGGCCCTCTGTCGGCGACGTTGATGTCAGCCCTC 282
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      22 Pro-----AsnAspPheAspValPheArgCysAsnGlnPheTyrPhe---GluAspLys 38
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      283 CCTATTAACTCGAGATTTT---GTATGTCATGCAATAACTTTTATGAATAAAGACCTT 339
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      39 TyrTyrLeuGlyLysCysLysTyrValPheTyrThrProAnpPhePheGluGln 58
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      340 AATACCGTCGGTGGATTTTCGCATCTATTGGCCCTACAGTTGAT----- 384
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      59 TyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCys 78
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
QY      385 -----TTATCATGAACATGCTCCAGGACCACATAGAAAACATGGAAGTGTTCG 435
Db      |||.....|||.....|||.....|||.....|||.....|||.....|||
79 SerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyr 98
436 GCT-----ATTGAAGGAGGATCTCTCTCTTTTATGAGACAAA 471
99 PheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyr 118
472 GAAGAGCTTATTCACCTTA-----AAAGAAATATCAAAAATAATATTGTT 516
119 PheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAla 138
517 TTT---TATCATCTCCGATATTTTCG---AACTTGGAAACATTCGGCAGATGTTTCT 570
139 ValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr----- 156
571 TTAGCAACTCACCTTCAAGTTTAAGAAATATTTTGTGGAATGCGCGGATATCCCGGA 630
157 -----GlnAsnGlySer 160
631 NAGCCCGCAATTGTTATGAACATGCCTTTGAAGGAAGTCAAAAGAACACCAAGGCGA 690
161 SerTyrAlaPheAspThrLysGlnGlu-----Asn 170
691 ACTTTTTCGTATGATTTCACAGGAAGGACCCCTGTNNCCCNNGATTAATTATGAAC 750
171 LeuLeuLysLeuAlaProAspPheLysAsn---AspArgSerHisTyrIleGlyHisSer 189
751 ACCTTAGTTTCCAAGGTAAATATATCAGAAATTTAGGAGAAGGCCACGACCAATCACTCG 810
190 LysAsnThrAspIleLysAla-LeuGluPheLeuGluLysThrTyrLysIleLysLeuTy 209
811 -----ACTGATATTTCTAGACCATATTATTTCTTTGAAATAATCCCAATAATATGTTAT 864
209 rCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSe 229
865 T-----AACTTTAGTTTATTAGGAGATTATGACCATTCCTCTCTCTATTCGCA 912
229 rAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGI 249
913 AAATCCGGTATTACCCCATCG---GGTATATCCGAAAGAACCAATCTTTTATTCGCCGC 969
249 uAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysLys 262
970 CCCCACGGGGGGGGGCAAGCCGTTTAAATTTCCAAAG 1009
RESULT 7
CL024751/c 1130 bp DNA linear GSS 31-DEC-2003
LOCUS CH216-20C18_RM1.1 CH216 Xenopus tropicalis genomic clone
DEFINITION CH216-20C18, genomic survey sequence.
ACCESSION CL024751 GI:40466756
VERSION CL024751
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Anura; Mesobatrachia; Pipidae;
Xenopodidae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1130)
AUTHORS Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Willson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM1_TACGACTCATATAGGAGA
Class: BAC ends
High quality sequence start: 117
High quality sequence stop: 714.
Location/Qualifiers
1..1130
FEATURES             source
```



```
Percent Similarity: 36.9%      Conservative: 52
Best Local Similarity: 21.3%    Mismatches: 137
Query Match: 6.9%              Indels: 74
DB: 6                          Gaps: 15

US-10-734-719-14 (1-291) x AY813556 (1-1619)

Qy 1 MetLysLysValIleIleAlaGlyAsn-----GlyProSerLeu 13
   |||||
Db 201 ATGAGAGAGAC-ATTATTTCAAGTAATGATGATGCATCTATATATGATGTTGAAGTGA 259
   |||||

Qy 14 LysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPhe 33
   |||||
Db 260 GAGTATATTGATTATCATGATCATCCATATGAAATGTTATTGATGATCTACTATTCTCT 319
   |||||

Qy 34 TyrPheGluAspLysTyrTyrLeuGlyLysCysLysThrValPheTyrThrProAsn 53
   |||||
Db 320 TATCAAAATGATGAATATGAAGAAGGTGAAGCGTATTTCACAAAGATTAT-----CCAAAT 373
   |||||

Qy 54 PhePhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThr 73
   |||||
Db 374 GTAGATTATCAAACTTATAATGAT-----GAGAAATGATCAATAT----- 412
   |||||

Qy 74 GluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLys 93
   |||||
Db 413 -----TATGGACCAAGTTCATATATATATATATGCAATATGATCAAA 451
   |||||

Qy 94 ThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLys 113
   |||||
Db 452 TCAAGTTATGATTAT-----AATATGCCATATATGATATATATTCACAAAGA 499
   |||||

Qy 114 GluPheAsnAlaTyrPheLysPheHis-----GluIleTyrPheAsnGlnArgIleThrSer 132
   |||||
Db 500 GATACAAATGATTATATCAAGAAAGACCTGATATACAATATCGTGTAGTCTTGGTAAT 559
   |||||

Qy 133 GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyr-----Leu 150
   |||||
Db 560 AGACCATATCAA-----TATCAAGAAGATTATGATGATGAT 595
   |||||

Qy 151 SerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsn 170
   |||||
Db 596 GACGCTGCTGATTACTATAGTAATGAAGGTAATTTATGAAGATGATTATACAGTAATAAT 655
   |||||

Qy 171 -----LeuLeuLysLeuAla 175
   |||||
Db 656 GATAATTATAAAGTCAATCTGATAGTAATCTAAATTTTCGCCAGTATTACAAATTC--- 712
   |||||

Qy 176 ProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThr----- 192
   |||||
Db 713 CCTACAGCATCATATAAACATCAAAAGTATGTACCGCATAGATTAGATGATTAAGAAAGAT 772
   |||||

Qy 193 -----AspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeu 208
   |||||
Db 773 TATCAACCAATTGATGATTAATATATATGATTTCTATCCAAAAGATATAAAACCAATAT 832
   |||||

Qy 209 TyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsn 228
   |||||
Db 833 TATACATTA-----CGTAAATATGCATACAAAATATATGTAAGCTAAATTCATATTTCTCA 886
   |||||

Qy 229 SerAsnPheIleIleGlnGluLys-----AsnAsnTyrThrLysAspIleLeuIleProSer 247
   |||||
Db 887 TCAAGATAATATCAAAAAATCACCAAATGGTACTTTTAAATATGTTTACCAACTACA 946
   |||||

Qy 248 -----SerGluAlaTyrGlyLysPhe-----SerLysAsnIle 258
   |||||
Db 947 TATCAATCTGATCATCATATAACATAGAGATAATTTCTATGATGATCAACAAAAATAAT 1006
   |||||

Qy 259 AsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 278
   |||||
Db 1007 CATTTATTACAAATCAATAATAAATCACTTAATATATAATCGATTGATTAATATTCA 1066
   |||||

Qy 279 ArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
   |||||
```

```
Db 1067 CCAACTAGCAACTATAAAAGATCATCATTTTAAATGGTAGA 1105

RESULT 9
DV791232/c
LOCUS
DEFINITION Hw_liver_21_050715_B11 Bos taurus CF-24-HW liver cDNA library Bos
            taurus cDNA, mRNA sequence.
ACCESSION DV791232
VERSION DV791232.1 GI:82644127
KEYWORDS EST.
SOURCE Bos taurus (cattle)
ORGANISM Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
          Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1376)
AUTHORS Yoon,D.H., Lee,S.H., Park,E.W., Cho,Y.M., Lee,J.H., Kim,H.,
          Kim,H.Y., Park,J.H. and Oh,S.J.
TITLE Gene Expression Profiling of the Bovine liver, adipose, and
          skeletal muscle
JOURNAL Unpublished (2005)
COMMENT Contact: Dr. Du-Hak Yoon
          National Livestock Research Institute, RDA
          564 Omockchun-dong, Suwon, 441-350, Korea
          Tel: 82 31 290 1593
          Fax: 82 31 290 1792
          Email: dhyoon@rda.go.kr.
FEATURES             location/Qualifiers
             source
               1..1376
               /organism="Bos taurus"
               /mol_type="mRNA"
               /db_xref="taxon:9913"
               /sex="Four males mixed"
               /cell_type="hepatocyte"
               /dev_stage="24 months old"
               /lab_host="XLI-BlueMRF, strain"
               /clone_lib="Bos taurus CF-24-HW liver cDNA library"
               /note="Organ: liver tissue; Vector: Uni-ZAPXR; Site_1:
               EcoRI; Site_2: Xho I"

ORIGIN
Alignment Scores:
Pred No.: 0.64      Length: 1376
Score: 107.50      Matches: 74
Percent Similarity: 32.3%      Conservative: 36
Best Local Similarity: 21.7%      Mismatches: 106
Query Match: 6.9%      Indels: 125
DB: 10              Gaps: 16

US-10-734-719-14 (1-291) x DV791232 (1-1376)

Qy 28 PheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysThr 47
   |||||
Db 1273 TTCCACACACACCAACACCACTCAAAACCAACCTACATATATCAAAACATTTATTTCTTT 1214
   |||||

Qy 48 ValPheTyrThrPro-----AsnPhePheGlu 57
   |||||
Db 1213 TTTTTCCTCCCAACCAACCCCAACCAACCAACCAACCAATTTTATACAAATTTTTCACAAAC 1154
   |||||

Qy 58 GlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMet 77
   |||||
Db 1153 CCCCATATCCACACATATACCATATATACATAAACCTTAT-----ATTTC 1106
   |||||

Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPhe----- 91
   |||||
Db 1105 TATATTAATATTATAAAATTCACACTAAAAAATAATATATCTATTTTTCAAAAATA 1046
   |||||

Qy 92 -----ValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhe 109
   |||||
Db 1045 CATTCACACACAATTTTCTTTTATTTTCCCCCTAAATAATA-----AAATTTT 992
   |||||

Qy 110 LysGlnLeuLysGluPhe-----Asn 116
   |||||
```





JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
DNA library

Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 17  
High quality sequence stop: 811.

## FEATURES

source

1..885

/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HM1:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHOsi; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

## ORIGIN

Alignment Scores:

Pred. No.: 0.407 Length: 885  
Score: 107.00 Matches: 67  
Percent Similarity: 34.4% Conservative: 31  
Best Local Similarity: 23.5% Mismatches: 92  
Query Match: 6.8% Indels: 97  
DB: 11 Gaps: 12

US-10-734-719-14 (1-291) x AZ529344 (1-885)

QY 21 LeuProAsnAspPheAspValPheAgcCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 679 TTGGCATCTCTTTCTGTTTATATTTCAATACACTTATATTGACGATGTTCCATAT 620  
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
DB 619 TATGGT-----ATTTTA 608  
QY 61 ThrLeuLysHisLeulleGlnAsnGlnGlu----- 70  
DB 607 ACAATATTATCATCTCTCTAACACACCAAGAAATTTGTAATTTCTACTGTGACTCAACAAGT 548  
QY 71 -----TyrGluThrGluLeulleMetCysSerAsnTyrAsnGlnAlaHisLeu 86  
DB 547 TCAATTGCACATATATTTATGAATATATTAATTTGTTGGNACAAATATGAATTAATATAT 488  
QY 87 GluAsnGluAsnPhe-----ValLysThrPheTyrAspTyrPheProAspAla 102  
DB 487 AAATCTCGTTCAGTGGTGGTGGCATCTATTCAAGATATAGTTTAGTGTGAAGCA 428  
QY 103 HisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHis 122  
DB 427 GATAAAATATTATTCATTTTATTTATTTGTTCAAAA-----GTTCAATCTCTCTTTTCA 377  
QY 123 GluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAla 142  
DB 376 TTATTATTATTCATCGTTGTTGTTTATTTTGGTAAGTTTAAATGTTTGTATTATTATAAG 317

QY 143 ---LeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer 161  
DB 316 TATTATCATATAAAGAAGAGATTAAACAATATACCTTATTTTCATATCAATCTCTC 257  
QY 162 TyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAsp 181  
DB 256 CCATTTTATCATGAAATGAATAAAT-----TCTATTAAACCTTCTGTTTAACGAAGAT 203  
QY 182 ArgSerHis-----TyrIleGlyHisSerLysAsnThrAsp 193  
DB 202 TGTATTATCATGATGAGTATGCTATTGCTATACATTAAAGGAGATGGAATCAATCAAC 143  
QY 194 IleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysPro 213  
DB 142 ATATCT-----TTATGT--- 131  
QY 214 AsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIle 233  
DB 130 -----AACTCTTTTAGTAATTTT----- 113  
QY 234 GlnGluLysAsnAsnTyrThrLysAspIleIleProSerSerGluAlaTyr---Gly 252  
DB 112 -----GATACCTTTAATAAC-AATAATGAAGCTACTGAAAC 78  
QY 253 LysPheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLys 272  
DB 77 AAATGAGT-----TATATCAAAA-AATGTTTATTACAAG 43  
QY 273 LeuIleLysAspLeu 277  
DB 42 ATTATAATGGGGTTG 28  
RESULT 13  
BM159659  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
BM159659 825 bp mRNA linear EST 04-DEC-2001  
EST562182 PyBS Plasmodium yoelii yoelii cDNA clone PYCJG93 5' end,  
mRNA sequence.  
BM159659 1 GI:17305340  
EST.  
Plasmodium yoelii yoelii  
Plasmodium yoelii yoelii  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
Carlton,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Vaidya,A.B.,  
Fraser,C.M. and Carucci,D.J.  
Plasmodium yoelii EST project at TIGR  
Unpublished (2001)  
Contact: Jane Carlton  
Parasite Genomics Group  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-530-9319  
Fax: 301-838-0208  
Email: carlton@tigr.org  
For clone info, please contact the Malaria Research and Reference  
Reagent Resource Center, ATCC  
http://www.malaria.mr4.org/mr4pages/index.html  
Seq primer: ADF.  
Location/Qualifiers  
1..825  
/organism="Plasmodium yoelii yoelii"  
/mol\_type="mRNA"  
/strain="17XL"  
/sub\_species="yoelii"  
/db\_xref="taxon:73239"  
/clone="PYCJG93"  
/dev\_stage="Asexual blood stages"  
/lab\_host="E. coli XL-1 Blue"  
/notes="Vector: pAD-GAL4; At 20-25% parasitemia, blood was  
collected from BALB/cByJ mice infected with Py17XL

parasites, and leukocytes removed by passage over microcrystalline cellulose columns. Total RNA was isolated using the guanidinium isothiocyanate method, and mRNA isolated using oligo(dT)-cellulose chromatography. First strand cDNA synthesis was completed using a 50-base primer and reverse transcriptase in the presence of 5-methyl dCTP. After second strand synthesis, uneven termini were treated with Pfu DNA polymerase and EcoRI adaptors ligated to the blunt ends. The sample was cleaved with XhoI and separated on a Sephacryl S-500 column. Size-fractionated cDNA was precipitated and ligated to HybriZAP arms directionally using EcoRI-XhoI cleaved arms. After packaging, the phagemid vector (pAD-GAL4) was excised from the HybriZAP vector and plasmid DNA isolated."

ORIGIN

Alignment Scores:  
Pred. No.: 0.48 Length: 825  
Score: 106.00 Matches: 55  
Percent Similarity: 39.2% Conservative: 38  
Best Local Similarity: 23.2% Mismatches: 86  
Query Match: 6.8% Indels: 58  
DB: 2 Gaps: 12

US-10-734-719-14 (1-291) x BM159659 (1-825)

Qy 27 ValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLys 46  
Db 95 ATATTAAATGTTAGTTTAAATAAACCCCAAAATATGCTATATATGAAAAAGAAAAAG 154  
Qy 47 ThrValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLe 66  
Db 155 TTAATTACATTTTCATGAAACATTTTAAATGAT-----ATA 190  
Qy 67 GlnAsnGlnGlyTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeu 86  
Db 191 GAAATAAAGAA-----ATTGGATCCCCAAATTTTATCATAC 229  
Qy 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106  
Db 230 ACTGGTTCAAAT-----TATGATTGGTTTTTAAATTC-----TTC 265  
Qy 107 AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPhe 126  
Db 266 GATTATGTTGGAGATACTAGTAAATAAAT-----AATTTAAGTTT 307  
Qy 127 AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys 146  
Db 308 AATAAATTTGTAATCT-----GTTGCAAAAGGACAGAAG 343  
Qy 147 GluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThr 166  
Db 344 AAA--AATAAGCATGGAATTTATTTTGCACAAT-----TTCATATTTGATAAT 391  
Qy 167 LysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIle 186  
Db 392 AAGAAGGATAACATAATA-----AATGAAAAAATAAAAACTAT 430  
Qy 187 GlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIle 206  
Db 431 TCACATATCGATTAAAGTAAGAAAGAAAAAATATTGTGGAAAAATAAAAAATATT 490  
Qy 207 LysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsn 226  
Db 491 ---TTATAT-----AAACGAACAGTGTGGATGTACTTCGAAC 526  
Qy 227 LeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuLeuPro 246  
Db 527 GTCAATACAAATCATACATACAGAAAAA-----ATAAATGCATATTTGATCAT 580  
Qy 247 SerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIle 263

Db 581 ATTAGTAATAACACTAATAAGAAATTAGTAAAAACATAAATAAGCAATATA 631  
RESULT 14  
CV987973  
LOCUS  
DEFINITION  
CV987973 730 bp mRNA linear EST 01-DEC-2004  
IPCGFr1\_10\_G08\_23 IPCGFr1 Ictalurus punctatus cDNA clone  
IPCGFr1\_10\_G08\_5', mRNA sequence.  
ACCESSION  
CV987973  
VERSION  
CV987973.1 GI:56163847  
KEYWORDS  
EST  
SOURCE  
ICталurus punctatus (channel catfish)  
ORGANISM  
ICталurus punctatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;  
Ictaluridae; Ictalurus.  
REFERENCE  
1 (bases 1 to 730)  
Li,R. and Waldbieser,G.  
Gene expression profiling by high-density oligonucleotide  
microarrays in the lipopolysaccharide-stimulated spleen of channel  
catfish  
Unpublished (2005)  
JOURNAL  
Contact: Waldbieser GC  
COMMENT  
Catfish Genetics Research Unit  
USDA-Agricultural Research Service  
141 Experiment Station Road, Stoneville, MS 38776, USA  
Tel: 662 686 3593  
Fax: 662 686 3567  
Email: gwaldbieser@ars.usda.gov  
Single pass sequencing. Bases called with Phred v0.000925.c. Low  
quality bases and vector trimmed with Lucy v1.16.  
Plate: 10 row: G column: 8  
Seq primer: SP6.  
FEATURES  
Location/Qualifiers  
1..730  
/organism="Ictalurus punctatus"  
/mol\_type="mRNA"  
/strain="NWAC103"  
/db\_xref="taxon:7998"  
/clone="IPCGFr1\_10\_G08"  
/dev\_stage="5-, 7-, and 9-day fry"  
/lab\_host="DH10B"  
/clone\_lib="IPCGFr1"  
/note="Organ: whole animal; Vector: pSport1; Site\_1: Sal  
1; Site\_2: Not 1; Primary library"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.465 Length: 730  
Score: 105.50 Matches: 39  
Percent Similarity: 40.8% Conservative: 30  
Best Local Similarity: 23.1% Mismatches: 73  
Query Match: 6.7% Indels: 27  
DB: 8 Gaps: 8  
US-10-734-719-14 (1-291) x CV987973 (1-730)  
Qy 30 CysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGly-----Lys 43  
Db 226 TGCACCAGCACCAACTCAAGCTTGAACATTATTCACACTTCTATAAAAGTCACCCAGGT 285  
Qy 44 LysCysLysThrValPheTyrThrProAsnPhe-----PhePheGluGlnTyrTyr 60  
Db 286 TCCTGTGCAAAACAACTCTTCAACAAACACTTCAACAACTCCACTTCAACAACTCCAC 345  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 346 TACACAAACACAACTTCAACAAACACCACTTCAACCCCGACGCTCTACAAACACGAC 405  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 406 TTCAACAACTCCACTTCAACAAACACCACTTCAACAAA-----CAACACTTCAGC 456  
Qy 101 AsplahisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyr---Phe 119

```

Db      457 AAACAACACTTCAGCAACACCACCTTCAACAG---CATGACTTCAATAAACAACGCTCTTC 513
Qy      120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyr-----MetCys 137
Db      514 AAACCTCAACCATCTAC-----AAACGGCTCTTCAACAAACTCTACTGTGCTAGCTGC 567
Qy      138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln 157
Db      568 AGTATTCAACTTGGTCTTTAGCATCAATGAAACCTTTTATCCAGCACTGGCGGATACCCAG 627
Qy      158 AsnGlySerSerTyrAlaPheAspThrLys-----Gln 168
Db      628 CTCCTCAGAGTTTACAGCAAAAGCTCAAAATATCCGTAATCAGCTTGAACCAATTTACAA 687
Qy      169 GluAsnLeuLeuLysLeuAlaProAsp 177
Db      688 ACAAAATCTTCNAAAACTTCCTCCGGAT 714

RESULT 15
AY411754
LOCUS      1131 bp      DNA      linear      GSS 16-DEC-2003
DEFINITION Pan troglodytes STAT8E gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION  AY411754
VERSION     AY411754.1  GI:39767722
KEYWORDS   Pan troglodytes (chimpanzee)
SOURCE      Pan troglodytes
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Pan.
REFERENCE   1 (bases 1 to 1131)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE   2 (bases 1 to 1131)
AUTHORS    Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT    This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES   Location/Qualifiers
source     1..1131
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            <1..>1131
            /gene="STAT8E"
            /locus_tag="HCM4329"

ORIGIN
Alignment Scores:
Pred. No.:      0.827      Length:      1131
Score:          105.50      Matches:      51
Percent Similarity: 40.0%      Conservative: 31
Best Local Similarity: 24.9%      Mismatches:  88
Query Match:    6.7%      Indels:      35
DB:              14      Gaps:        9

US-10-734-719-14 (1-291) x AY411754 (1-1131)

Qy      2 LysLysValIleAlaLeuGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21

```

```

Db      484 AAGAAGTGTGCTGTAGTGGGCAACGGAGGCATCTTTGAAGAACACGCCGTGTGGAGGGAG 543
Qy      22 ProAsnAspPheAsp---ValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      544 ATCAACAGCGCCGACTTTCGGTGCACCTGCCCCCATCTCAGAGAAGTACACC 603
Qy      41 Leu-----GlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGln 58
Db      604 ATGGAGCTGGGGGTGAAGCGGATGTGTCCTACTGTGAACCCACCATCATCACAGAGAGG 663
Qy      59 TyrTyrThrLeuLys-----HisLeuIleGlnAsnGlnGluTyr 71
Db      664 TTCCANNNGCTGAGAAAGTGGCGGCGCGCTTCTATCGCGTGTGTCAGGTGTACGAG--- 720
Qy      72 GluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPhe 91
Db      721 AACGCGTCCGCTGTGCTGCTGCTCTTCTACAACACGCGCAACACCGAGCTGTCATCCGCG 780
Qy      92 ValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGln 111
Db      781 GTCAAATACGTGCTGGACGACTTCGAATCGCCGCAAGCTGTCTACTACTTCCATCCGCGAG 840
Qy      112 ---LeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130
Db      841 TACCTGTCACACGTGTCGCGCTACTGGCTCAGCCCTGGGGGTGCGC---GCCAAGCGCATC 897
Qy      131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150
Db      898 AGCACCGGCTCATCTTCTGGTCACCTGGCGCGCTGGAGCTC---TGTGAGGAGGTGCACCTC 954
Qy      151 SerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsn 170
Db      955 TTTGGCTTC-----TGGGCTTCCCCATGAACCCCTCGGCGC 990
Qy      171 LeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLys 190
Db      991 CTC-----TACATCACTCACCACCTAC 1011
Qy      191 AsnThrAspIleLys 195
Db      1012 TATGACAACGTCAAG 1026

```

Search completed: August 21, 2006, 20:38:43  
Job time : 6314 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 15:00:51 ; Search time 176.5 Seconds

(without alignments)  
4627.416 Million cell updates/sec

Title: US-10-734-719-14

Perfect score: 1564

Sequence: 1 MKKVITAGNSPLKEIDYSR.....KLIKDLRLPLSDIKHYFKGK 291

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abss/ABSSWEB spool/US10734719/runat 21082006 141818 12415/app query.fasta\_1  
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosun62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs05p -USER=US10734719 @CGN 1 1 381 @runat 21082006 141818 12415  
-NCPU=6 -ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

1: /EMC Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq.\*  
2: /EMC Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq.\*  
3: /EMC Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq.\*  
4: /EMC Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq.\*  
5: /EMC Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq.\*  
6: /EMC Celerra\_SIDS3/ptodata/2/ina/H COMB.seq.\*  
7: /EMC Celerra\_SIDS3/ptodata/2/ina/PTUS COMB.seq.\*  
8: /EMC Celerra\_SIDS3/ptodata/2/ina/pp COMB.seq.\*  
9: /EMC Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq.\*  
10: /EMC Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1564	100.0	873	3	US-09-816-028A-13
2	1564	100.0	873	3	US-10-303-162-13
3	1564	100.0	873	3	US-10-303-134-13
4	1564	100.0	873	3	US-10-303-118-13
5	1564	100.0	873	3	US-10-303-128-13
6	1564	100.0	873	5	US-10-735-419-13
7	1559	99.7	876	3	US-09-495-406-8
8	1559	99.7	876	3	US-09-816-028A-8

9	1559	99.7	876	3	US-10-303-162-8	Sequence 8, Appli
10	1559	99.7	876	3	US-10-303-134-8	Sequence 8, Appli
11	1559	99.7	876	3	US-10-303-118-8	Sequence 8, Appli
12	1559	99.7	876	3	US-10-303-128-8	Sequence 8, Appli
13	1559	99.7	876	5	US-10-735-419-8	Sequence 8, Appli
14	1553	99.3	873	3	US-09-816-028A-11	Sequence 11, Appl
15	1553	99.3	873	3	US-10-303-162-11	Sequence 11, Appl
16	1553	99.3	873	3	US-10-303-134-11	Sequence 11, Appl
17	1553	99.3	873	3	US-10-303-118-11	Sequence 11, Appl
18	1553	99.3	873	3	US-10-303-128-11	Sequence 11, Appl
19	1553	99.3	873	5	US-10-735-419-11	Sequence 11, Appl
20	1524	97.4	876	3	US-09-495-406-4	Sequence 4, Appli
21	1524	97.4	876	3	US-09-816-028A-4	Sequence 4, Appli
22	1524	97.4	876	3	US-10-303-134-4	Sequence 4, Appli
23	1524	97.4	876	3	US-10-303-118-4	Sequence 4, Appli
24	1524	97.4	876	3	US-10-303-128-4	Sequence 4, Appli
25	1524	97.4	876	5	US-10-735-419-4	Sequence 4, Appli
26	1521	97.3	876	3	US-09-495-406-2	Sequence 2, Appli
27	1521	97.3	876	3	US-09-816-028A-2	Sequence 2, Appli
28	1521	97.3	876	3	US-10-303-162-2	Sequence 2, Appli
29	1521	97.3	876	3	US-10-303-134-2	Sequence 2, Appli
30	1521	97.3	876	3	US-10-303-118-2	Sequence 2, Appli
31	1521	97.3	876	3	US-10-303-128-2	Sequence 2, Appli
32	1521	97.3	876	5	US-10-735-419-2	Sequence 2, Appli
33	1521	97.3	876	3	US-09-495-406-1	Sequence 1, Appli
34	1521	97.3	11474	3	US-09-816-028A-1	Sequence 1, Appli
35	1521	97.3	11474	3	US-10-303-162-1	Sequence 1, Appli
36	1521	97.3	11474	3	US-10-303-134-1	Sequence 1, Appli
37	1521	97.3	11474	3	US-10-303-118-1	Sequence 1, Appli
38	1521	97.3	11474	3	US-10-303-128-1	Sequence 1, Appli
39	1521	97.3	11474	5	US-10-735-419-1	Sequence 1, Appli
40	1521	97.3	11474	5	US-09-495-406-6	Sequence 6, Appli
41	1510	96.5	876	3	US-09-816-028A-6	Sequence 6, Appli
42	1510	96.5	876	3	US-10-303-162-6	Sequence 6, Appli
43	1510	96.5	876	3	US-10-303-134-6	Sequence 6, Appli
44	1510	96.5	876	3	US-10-303-118-6	Sequence 6, Appli
45	1510	96.5	876	3	US-10-303-128-6	Sequence 6, Appli

#### ALIGNMENTS

RESULT 1

US-09-816-028A-13

; Sequence 13, Application US/09816028A

; Patent No. 6699705

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-000111US

; CURRENT APPLICATION NUMBER: US/09/816,028A

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/118,213

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406

; PRIOR FILING DATE: 2000-01-31

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 873

; TYPE: DNA

; ORGANISM: Campylobacter jejuni

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(873)

; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
(CstII) from C. jejuni O:36

US-09-816-028A-13

Alignment Scores:

Pred. No.: 5.84e-173 Length: 873

```
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-09-816-028A-13 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTATTATTCGTGAAATGGCAAGTTTAAAGAAATTTGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheAtqCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGATGATTAGATGAATCAATTTTATTTTGAAGATAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTTAATCCAAAATCAAGAAATATGAGACCGCACTAATTTATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAAACAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTACGCAAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTACAAAAATAAAACATATATGCTTATGCTCAATATAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAAAATAAGAAAAATGTTTATTAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873

RESULT 2
US-10-303-162-13
; Sequence 13, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-162-13

Alignment Scores:
Pred. No.: 5,84e-173 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-162-13 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTATTATTCGTGAAATGGCAAGTTTAAAGAAATTTGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGATGATTAGATGAATCAATTTTATTTTGAAGATAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTTAATCCAAAATCAAGAAATATGAGACCGCACTAATTTATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAAACAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTACGCAAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTACAAAAATAAAACATATATGCTTATGCTCAATATAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAAAATAAGAAAAATGTTTATTAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873
```

RESULT 2

US-10-303-162-13

; Sequence 13, Application US/10303162

; Patent No. 6723545

; GENERAL INFORMATION:

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
| | | | |  
Db 541 GATCGCTCACACTATATCGCACATAGTAAATAACAGATATAAAGCTTTAGATTCTTA 600  
| | | | |  
QY 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
| | | | |  
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTAGCAAAATTTT 660  
| | | | |  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
| | | | |  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACACT 720  
| | | | |  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
| | | | |  
Db 721 AAAGATATACTATACCTTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
| | | | |  
QY 261 LysLysIleLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
| | | | |  
Db 781 AAAAAAATAAAAAATAAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
| | | | |  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
| | | | |  
Db 841 CCTAGTGATATAAAGCATTTATTTCAAGGAAAA 873  
| | | | |

## RESULT 3

US-10-303-134-13  
; Sequence 13, Application US/103031134  
; Patent No. 6825019  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-134-13

## Alignment Scores:

Pred. No.: 5.84e-173 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-134-13 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
| | | | |  
Db 1 ATGAAAAAAGTTTATTATGCTGGAAATGGCAAGTTTAAAGAAAAATTTGATTATTCAGG 60  
| | | | |  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
| | | | |  
Db 61 CTACCAATGATTTTGATGATTAGATGTAATCAATTTTATTATTGTAAGATAAATACTAT 120  
| | | | |

QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60  
| | | | |  
Db 121 CTTGGTAAAAATGCAAAAACAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180  
| | | | |  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
| | | | |  
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATATATGTGTCTTAAT 240  
| | | | |  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
| | | | |  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTATTTCCT 300  
| | | | |  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
| | | | |  
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAATTTAATGCTTATTATAA 360  
| | | | |  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
| | | | |  
Db 361 TTTCACGAAATTTATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
| | | | |  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrClnAsnGlySer 160  
| | | | |  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480  
| | | | |  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
| | | | |  
Db 481 TCATTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540  
| | | | |  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
| | | | |  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
| | | | |  
QY 201 GluLysThrTyrLysLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
| | | | |  
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660  
| | | | |  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
| | | | |  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACACT 720  
| | | | |  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
| | | | |  
Db 721 AAAGATATACTATACCTTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
| | | | |  
QY 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
| | | | |  
Db 781 AAAAAAATAAAAAATAAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
| | | | |  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
| | | | |  
Db 841 CCTAGTGATATAAAGCATTTATTTCAAGGAAAA 873  
| | | | |

## RESULT 4

US-10-303-118-13  
; Sequence 13, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-118-13

Alignment Scores:
Pred. No.: 5,84e-173 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-118-13 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTTTATGATGTAATCAATCAATTTTATTGTAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATGCAAAACAGTGTATACACCCCTAAATTTCTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATTATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCAGCAATTTATTTCAATCAAGAATTACCTCAGGGGCTATATGTCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTGATACCAACCAAGAAAATCTTTTAAAACTAGCCCTCATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACATATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATACAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATCAAGAAAAAATAACTACACT 720
QY 241 LysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTTCTAGTAGGGCTATGGAAAAATTTTCAAAAAATATTAATTTT 780

; 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
; 781 AAAAAAATAAATAAAGAAATGTTTATTACAGTTGATAAAGATCTATTAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAGCATTATTTCAAAAGGAAAA 873

RESULT 5
US-10-303-128-13
; Sequence 13, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-128-13

Alignment Scores:
Pred. No.: 5,84e-173 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-128-13 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTTTATGATGTAATCAATCAATTTTATTGTAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATGCAAAACAGTGTATACACCCCTAAATTTCTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATTATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCAGCAATTTATTTCAATCAAGAATTACCTCAGGGGCTATATGTCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTGATACCAACCAAGAAAATCTTTTAAAACTAGCCCTCATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACATATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATACAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATCAAGAAAAAATAACTACACT 720
QY 241 LysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTTCTAGTAGGGCTATGGAAAAATTTTCAAAAAATATTAATTTT 780
```

Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnAtqIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTATATATGTGCGAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAACTTACAAATAAACATATATGCTTAATGCTCTTAATAGCTTTTAGCAAAATTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAATTTTCAAAATTTT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTCTTAGTAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAAAATTAAGAAATGTTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873  
RESULT 6  
US-10-735-419-13  
; Sequence 13, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-735-419-13  
Alignment Scores: 5.84e-173 Length: 873  
Pred. No.: 1564.00 Matches: 291  
Score:

Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 5 Gaps: 0  
US-10-734-719-14 (1-291) x US-10-735-419-13 (1-873)  
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTTATTTGCTGGAATCGAACAGTTTAAAAAGAAATTCATTATTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCNAATCTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCGCACTAATATTATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAACTTACAAATAAACATATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAATTTTCAAAATTTT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTCTTAGTAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAAAATTAAGAAATGTTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873  
RESULT 7  
US-09-495-406-8  
; Sequence 8, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel

```
/ APPLICANT: Wakarchuk, Warren W.
/ APPLICANT: National Research Council of Canada
/ TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
/ TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
/ FILE REFERENCE: 019633-000110US
/ CURRENT APPLICATION NUMBER: US/09/495,406
/ CURRENT FILING DATE: 2000-01-31
/ PRIOR APPLICATION NUMBER: US 60/118,213
/ PRIOR FILING DATE: 1999-02-01
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 876
/ TYPE: DNA
/ ORGANISM: Campylobacter jejuni
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(876)
/ OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
/ OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-495-406-8

Alignment Scores:
Pred. No.: 2,25e-172 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-09-495-406-8 (1-876)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGTGGAATGACCAAGTTTAAAGAAAATTCATTATTCACAG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTTGATGATTAATAGATGAATCAATTTTATTTGAAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAACAGCGTTTACACCCCTAAATTCCTCTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAAAACATTAATCCAAAATCAAGAAATATGAGACCGAACTAATAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTACCATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAGAAATTAATGCTTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACAGCTACCCCTGATTTTAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTGCACCTATATCGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
```

```
DB 601 GAAAAAATTACAAAATAAACTATATTGCTTATGCTTAATAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAATTAAGAAAATGTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 8
US-09-816-028A-8
/ Sequence 8, Application US/09816028A
/ Patent No. 6699705
/ GENERAL INFORMATION:
/ APPLICANT: Gilbert, Michel
/ APPLICANT: Wakarchuk, Warren W.
/ APPLICANT: National Research Council of Canada
/ TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
/ TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
/ FILE REFERENCE: 019633-000110US
/ CURRENT APPLICATION NUMBER: US/09/816,028A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US 60/118,213
/ PRIOR FILING DATE: 1999-02-01
/ PRIOR APPLICATION NUMBER: US 09/495,406
/ PRIOR FILING DATE: 2000-01-31
/ NUMBER OF SEQ ID NOS: 49
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 876
/ TYPE: DNA
/ ORGANISM: Campylobacter jejuni
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)..(876)
/ OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
/ OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-816-028A-8
```

```
Alignment Scores:
Pred. No.: 2,25e-172 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0
```

```
US-10-734-719-14 (1-291) x US-09-816-028A-8 (1-876)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGTGGAATGACCAAGTTTAAAGAAAATTCATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTTGATGATTAATAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
```

```
Db 181 ACTTTAAACATTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATATGTGTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAAACTTAAGAAATTAATGCTTATTTTAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTGCTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATTAACATATATGCTTATGTCTCTATAGTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAAGAAATAATCACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTTCTAGTGAGCTTATCGAAAATTTTCAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTAATTTTCAAGGAAAA 873
```

## RESULT 9

```
US-10-303-162-8
; Sequence 8, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-162-8
```

## Alignment Scores:

```
Pred. No.: 2,25e-172 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-14 (1-291) x US-10-303-162-8 (1-876)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTTATTTCTGGAAATGACCAAGCTTTAAAAAGAAATTCGATTATTCACGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGTATTTTGAIGTATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTTCGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGCCGAATAATATATGTGTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAAACTTAAGAAATTTAATGCTTATTTTAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTGCTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATAACAGATATAAAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATTAACATATATGCTTATGTCTCTAATAGTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTCAAAATTTTATCATACAAAGAAATAAATACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTTCTAGTGAGCTTTATGAAAAATTTTCAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
```





Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTGATGTTAGTATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGlyThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAAAATTTTAAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTTATTTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCAATTTGGGATATGATTTTTTAAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnAtgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGATGACC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCGACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAATAACTATATTGCTTATGTCCTATAGCTTTTATGCAAAATTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnLysAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATCCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTTAAATTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAATAAATAAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTTCAAGGAAAA 873

RESULT 12

US-10-303-128-8  
; Sequence 8, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-128-8

Alignment Scores:

Pred. No.: 2,25e-172 Length: 876  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-128-8 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTTATGATGTTAGTATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyrTyr 60  
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGlyThrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAAAATTTTAAATCCAAATCAAGAAATAGACCGAACTAATATGTTCTTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTTATTTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCAATTTGGGATATGATTTTTTAAAACAACTTAAAGAAATTTTAAATGCTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGATGACC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGATACAAAGAAATTTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCGACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAATAACTATATTGCTTATGTCCTATAGCTTTTATGCAAAATTT 660

```
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAATAAATAACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTAGTAGGCTTATGGAAATTTTCAAAAAATATAATTTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATT 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 13
US-10-735-419-8
; Sequence 8, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-735-419-8

Alignment Scores:
Pred. No.: 2,25e-172 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 5 Gaps: 0

US-10-734-719-14 (1-291) x US-10-735-419-8 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGACGATGTTTATACACCCCTAAATTTCTTTGTAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAAAACATTAAATCCAAAAATCAAGAAATATGAGACCGAAGTAATTTATGTGTTCTAAT 240
```

```
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAAACAGCTCATCTAGAAAATGAAAATTTTGTGAAAACTTTTACAGTATTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAITTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAAAATAAACTATATGTCTTATGCTTATGCTTATGCTTTTACGAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATCAAGAAAAATAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATAATTTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATT 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 14
US-09-816-028A-11
; Sequence 11, Application US/09816028A
; Patent No. 6695705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
```

## US-09-816-028A-11

Alignment Scores:  
Pred. No.: 1,12e-171 Length: 873  
Score: 1553.00 Matches: 289  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 2  
Query Match: 99.3% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-09-816-028A-11 (1-873)

```
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
    |||
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTCAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
    |||
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
    |||
Db 121 CTTGGTAAAAAATGCAAGCAGTCTTTTACACCCCTGGTTTCTTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
    |||
Db 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
    |||
Db 241 TACCAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTTCT 300

Qy 101 AspalatHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
    |||
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
    |||
Db 361 TTTTACGAAATTTATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
    |||
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
    |||
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
    |||
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
    |||
Db 601 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTACAGTCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
    |||
Db 661 ATAGAACTAGCGCAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
    |||
Db 721 AAAGATATCTCATACCTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTTAATTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
    |||
Db 781 AAAAAATAAAAAATAAGAAAAATGTTTATACAAAGTTTGATAAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
    |||
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
```

RESULT 15

## US-10-303-162-11

; Sequence 11, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4

## US-10-303-162-11

Alignment Scores:  
Pred. No.: 1,12e-171 Length: 873  
Score: 1553.00 Matches: 289  
Percent Similarity: 99.3% Conservative: 0  
Best Local Similarity: 99.3% Mismatches: 2  
Query Match: 99.3% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-162-11 (1-873)

```
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
    |||
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
    |||
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
    |||
Db 121 CTTGGTAAAAAATGCAAGCAGTCTTTTACACCCCTGGTTTCTTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
    |||
Db 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
    |||
Db 241 TACCAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTTCT 300

Qy 101 AspalatHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
    |||
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
    |||
Db 361 TTTTACGAAATTTATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
    |||
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
```



GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 15:32:54 ; Search time 1385 Seconds  
(without alignments)  
3872.601 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIIAGNPSLKEIDYSR.....KLIKDLRLPDSIKHYFKGK 291

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 18992170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abss/ABSSWEB spool/US10734719/runat.21082006.141829.12577/app\_query.fasta\_1  
-DB=Published Applications NA Main -QFIX=fastap -SUFFIX=P2N.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LCOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pto -NORM=ext  
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss03h  
-USER=US10734719 @CCN 1.1 2326 @runat.21082006.141829.12577 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:\*  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	1564	100.0	873 3 US-09-816-028A-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-816-028A-13  
; Sequence 13, Application US/09816028A  
; Patent No. US20020042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816.028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-09-816-028A-13

Alignment Scores:  
Pred. No.: 6,65e-150 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-14 (1-291) x US-09-816-028A-13 (1-873)

Qy 1 MetLysValIlelelaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTCGCAATGACCAAGTTTAAAGAAATGATTATTCAAG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAAAATGATTTTGAATGATTAATCAATTTTATTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAAAAATGCAAAACAGTGTTTTACCCCTTAATTTCTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluThrGluLeuMetCysSerAsn 80  
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTAATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAAAATTTTACGATTTATTTCTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAGAAATTTAATGCTTATTTTAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTAATTCATCAAGAAATTAACCTACCTCAGGGTCTATATGTCAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAGAAATTCCTTTAAAACTAGCCCTCGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAAATAAATATATATGCTATGTCCTAATAGTCTTTTAGCAAAATTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCCCAAAATTTAAATTCAAATTTTATCATCAAGAAAAATAAATACTACCT 720  
Qy 241 LysAspIleLeuIleProSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATTTTCAAAAAATATATATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAATAAAGAAATGTTTATTAACAGTTGATAAAGATCTATTAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291

Db 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

## RESULT 2

US-10-303-161-13  
; Sequence 13, Application US/10303161  
; Publication No. US20030148459A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,161  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 05/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-161-13

Alignment Scores:  
Pred. No.: 6,65e-150 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 7 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-161-13 (1-873)

Qy 1 MetLysValIlelelaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTCGCAATGACCAAGTTTAAAGAAATGATTATTCAAG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTTGAATGATTAATCAATTTTATTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAAAAATGCAAAACAGTGTTTTACCCCTTAATTTCTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluThrGluLeuMetCysSerAsn 80  
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAGAAATTTAATGCTTATTTTAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTTATTCATCAAGAAATTAACCTACCTCAGGGTCTATATGTCAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTATATGCTTTTGATACCAACCAAGAAATCTTTTAAAGCTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGATTTCTA 600  
Qy 201 GluLysThrTyrLysLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 220  
Db 601 GAAAAAATTTACAAATAAAATATATATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATCAAGAAAGAAATTAACACACT 720  
Qy 241 LysAspIleLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 260  
Db 721 AAAGATATATCTATACCTCTAGTAGGCTTTATGGAATTTTCAAAAAATATTAATTTT 780  
Qy 261 LysLysIleLysLysLeuLysGluAsnValTyrTyrLysLeuLysLeuLysLeuLysLeu 280  
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTCAAGGAAAA 873

## RESULT 3

US-10-303-118-13  
; Sequence 13, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-118-13

## Alignment Scores:

Pred. No.:	6,65e-150	Length:	873
Score:	1564.00	Matches:	291
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	7	Gaps:	0

US-10-734-719-14 (1-291) x US-10-303-118-13 (1-873)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTTATTTCTGGAAATGGACCAAGTTTAAAAAGAAATTCGATTATTCAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTGTATGTTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTCTTTGGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATATGTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTTCCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACACTTAAAGAAATTTAATGCTTATTTTAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTGTCAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrIleuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATCAAGAAAGAAATTAACACACT 720  
Qy 241 LysAspIleLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 260  
Db 721 AAAGATATATCTATACCTCTAGTAGGCTTTATGGAATTTTCAAAAAATATTAATTTT 780  
Qy 261 LysLysIleLysLysLeuLysGluAsnValTyrTyrLysLeuLysLeuLysLeuLysLeu 280  
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTCAAGGAAAA 873

## RESULT 4

US-10-303-128-13  
; Sequence 13, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21

```
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-128-13

Alignment Scores:
Pred. No.: 6,65e-150 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-128-13 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATCTGCTGGAATGACCAAGTTTAAAGAAATTGATTATTCACAGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTTGATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAAGTAAATATGTTGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAAAAATTTTACCGATTATTTTCTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTATTTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGAGTAGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAAACTAGCCCTCATTTTAAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTCAAAAATAAAATATATGCTTATGTTCTTATAGTCTTTTAGCAAAATTTT 660
```

```
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAATTTTATCATACAGAAAAATACTACTACT 720

QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTTCTAGTGAGCTTATGAAAAATTTTCAAAAAATATTAAATTTT 780

QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAAAGATCTATTAGATTA 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 5
US-10-303-134-13
; Sequence 13, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-134-13

Alignment Scores:
Pred. No.: 6,65e-150 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-734-719-14 (1-291) x US-10-303-134-13 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTTGCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCACAGG 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTTGATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAAGTAAATATGTTGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAAAAATTTTACCGATTATTTTCTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTATTTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGAGTAGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAAACTAGCCCTCATTTTAAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTCAAAAATAAAATATATGCTTATGTTCTTATAGTCTTTTAGCAAAATTTT 240
```







```
QY 1 MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCAGAACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTTATTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACAGAAATTTATTTCAATCAAGAAATACCTTCAGGGGTCTATATGTGTGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAATATAAATCTATATGCTTAATGCTCTAATAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAATAAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATT 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTATATAAGCATTTATTTCAAAGGAAAA 873
```

## RESULT 9

```
US-10-845-408-13
; Sequence 13, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
```

```
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-408-13
```

## Alignment Scores:

Pred. No.:	6.65e-150	Length:	873
Score:	1564.00	Matches:	291
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	9	Gaps:	0

US-10-734-719-14 (1-291) x US-10-845-408-13 (1-873)

```
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCAGAACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTTATTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACAGAAATTTATTTCAATCAAGAAATTTTCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAATATAAATCTATATGCTTAATGCTCTAATAGTCTTTTAGCAAAATTT 660
```

```
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleileGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTCTTATCATCAAGAAATAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTCAAAAATATTAATTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATACAGTTGATAAAAGACTATTAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 10
US-10-845-412-13
; Sequence 13, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; PRIOR FILING DATE: 2004-05-12
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-412-13

Alignment Scores:
Pred. No.: 6,65e-150 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-734-719-14 (1-291) x US-10-845-412-13 (1-873)

QY 1 MetLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAAAATTCATTATCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTCATGATTAATGATTAATCAATTTTATTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTTTGGCAATACTAC 180
```

```
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAAACATTTAATCCAAAATCAAGAATATGAGACCGAAGCTAATATATGTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAATAAACTTTTACGATTATTTTCCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCAGTAGGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTATGATCAAAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATTCGTTATGCTCTAATAGTCTCTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleileGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTTAAATTTCAATTTTATCATCAAGAAAAATAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAATATTAATTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 11
US-10-846-219-13
; Sequence 13, Application US/10846219
; Publication No. US20040219638A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
```

; FEATURE: CDS  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
; US-10-846-219-13

## Alignment Scores:

Pred. No.: 6.65e-150 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 9 Gaps: 0

US-10-734-719-14 (1-291) x US-10-846-219-13 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGCTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACATC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATATGTTCTTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAATAAACTTTTACGATTAATTTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGTCAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTGATACCAACAAAGAAATCTTTTAAACAACTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTAGCAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGGCAAAATTTAAATTTCAATTTTATCATCAAGAAAAAATACTACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACCTACCTCTAGGAGCTTATGGAATAATTTTCAAAAAATATTTAAATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAATAAATAAAGAAATGTTTATTTATCAAGTTTGATAAAGATCTATTAAATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAAGCATTTATTTCAAGGAAAA 873

## RESULT 12

US-10-821-604-13  
; Sequence 13, Application US/10821604  
; Publication No. US2004029263A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-821-604-13

Alignment Scores:  
Pred. No.: 6.65e-150 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 9 Gaps: 0

US-10-734-719-14 (1-291) x US-10-821-604-13 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTATTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACATC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATATGTTCTTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAATAAACTTTTACGATTAATTTCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

Db 361 TTTTCAGGAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLysLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATAAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTATCATCAAGAAAAATAACTACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACTCTTAGTGAGGCTTAGGAAATTTTCAAAAATATATTAATTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

## RESULT 13

US-10-847-983-13  
; Sequence 13, Application US/10847983  
; Publication No. US20040229272A1  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Michel  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/847,983  
; CURRENT FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-847-983-13

Alignment Scores:  
Pred. No.: 6,65e-150 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0

Db: 9 Gaps: 0  
US-10-734-719-14 (1-291) x US-10-847-983-13 (1-873)  
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAAGCTTATTTATGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60  
QY 21 LeuProAsnAspPheAspValPheAtcCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAAATGCAAAAACAGTGTTTTACACCCCTAATTTCTCTTTTGACCAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGACCGGAACCTAATTTATGTGTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAACGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATTTTTCCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAGAAATTTAATGCTTTATTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGGAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATAAAACTATATGCTTATGTCCTTATAGTCTTTTACCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTATCATCAAGAAAAATAACTACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACTCTTAGTGAGGCTTAGGAAATTTTCAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

## RESULT 14

US-10-821-573-13  
; Sequence 13, Application US/10821573  
; Publication No. US20040229313A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

```
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-821-573-13

Alignment Scores:
Pred. No.: 6.65e-150 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-734-719-14 (1-291) x US-10-821-573-13 (1-873)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCGCGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTCGTAATAAATGCAAAACAGTGTTCACACCCCTAATTTCTCTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTATTTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGluLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTTAATGCTTATTTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCATCAAAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 941 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

541 GATCGCTCACACTATATCGGACATAGTAAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
601 GAAAAAACTTACAAAATAAAACTATATTCCTTATGCTCTAATAGTCTTTTACCAATTTT 660
221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
661 ATAGAAGTAGCGCCAAATTTAAATTCAAATTTTATCATACACAAAAAATAACTACTACT 720
241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
721 AAAGATATCTCATACCTTCTAGTGAGCTTATGAAAAATTTCAAAAAATTTAATTTT 780
261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAGATTATTA 840
281 ProSerAspIleLysHisTyrPheLysGlyLys 291
841 CCTAGTATATAAGCATTATTTTCAAGGAAAA 873

RESULT 15
US-10-850-807-13
; Sequence 13, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-850-807-13

Alignment Scores:
Pred. No.: 6.65e-150 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-734-719-14 (1-291) x US-10-850-807-13 (1-873)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCGCGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTCGTAATAAATGCAAAACAGTGTTCACACCCCTAATTTCTCTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTATTTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGluLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTTAATGCTTATTTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCATCAAAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 941 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
```

121	CTTGGTAAAAAATCAAAACAGTGTCTTTACACCCCTAATTTCTCTTTGAGCAATACTAC	180
61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
181	ACTTTAAACAACTTAATCCAAAATCAAGAAATAGAGCCGAACCTAATATGTGTTCTAAT	240
81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATATATTTCTCT	300
101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	360
121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
361	TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC	420
141	IleAlaIleuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
421	ATAGCCCTAGGATCAAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA	480
161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT	540
181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
541	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
201	GluLysThrTyrLysIleLysLeuTyrCysIeuCysProAsnSerLeuLeuAlaAsnPhe	220
601	GAATAAACTTTACAAAAATAAAACTATATTGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT	660
221	IleGluIleuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
661	ATAGAACTAGCGCCCAATTTTAAATTTCAATTTATCATACAGAAAAAAATAACTACACT	720
241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
721	AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTTAATTTT	780
261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
781	AAAAAATAAAAAATAAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAGATTTA	840
281	ProSerAspIleLysHisTyrPheLysGlyLys	291
841	CCTAGTCATATAAGCATTTATTTCAAGGAAAA	873

Search completed: August 21, 2006, 16:48:30  
Job time : 1391 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 15:36:45 ; Search time 195.5 Seconds  
3569.177 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIAGNPSLKEIDYSR.....KLIKDLRLUPSDIKHYFKGK 291

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2226170 seqs, 799283156 residues

Total number of hits satisfying chosen parameters: 4452340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DBV=xlh  
-Q=/abs/ABSSWEB.spool/US10734719/runat\_21082006\_141832\_12628/app\_query.fasta.1  
-DB=Published Applications NA.New -QWMT=fastap -SUFFIX=P2N.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MTN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -HOST=abes04  
-USER=US10734719 @CGN 1.1 286 @runat\_21082006\_141832\_12628 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :				Published Applications NA.New:*			
1:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US09	NEW_PUB.seq:*	Sequence 13, Appl	US-10-734-719-13	7.2	1659
2:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US06	NEW_PUB.seq:*	Sequence 8, Appl	US-10-734-719-8	7.1	11474
3:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US07	NEW_PUB.seq:*	Sequence 11, Appl	US-10-734-719-11	6.8	1614
4:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US08	NEW_PUB.seq:*	Sequence 4, Appl	US-10-734-719-4	6.7	2103
5:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US10	NEW_PUB.seq:*	Sequence 2, Appl	US-10-734-719-2	6.6	7500
6:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US11	NEW_PUB.seq:*	Sequence 1, Appl	US-10-734-719-1	6.6	1953
7:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US11	NEW_PUB.seq1.*	Sequence 1, Appl	US-10-734-719-1	6.5	1248
8:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US11	NEW_PUB.seq2.*	Sequence 1, Appl	US-10-734-719-1	6.5	109669
9:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US11	NEW_PUB.seq2.*	Sequence 1, Appl	US-10-734-719-1	6.5	2175
10:	/EMC_Celerra	SIDS3/ptodata/2/pubpna/US60	NEW_PUB.seq.*	Sequence 1, Appl	US-10-734-719-1	6.5	2931
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.				SUMMARIES			
Result No.	Score	Query Match	Length DB ID	Description			
1	1564	100.0	873	6	US-10-734-719-13	7.2	1659
2	1559	99.7	876	6	US-10-734-719-8	7.1	11474
3	1553	99.3	873	6	US-10-734-719-11	6.8	1614
4	1524	97.4	876	6	US-10-734-719-4	6.7	2103
5	1521	97.3	876	6	US-10-734-719-2	6.6	7500
6	1521	97.3	11474	6	US-10-734-719-1	6.6	1953
7	1510	96.5	876	6	US-10-734-719-6	6.5	1248

ALIGNMENTS

RESULT 1  
US-10-734-719-13  
; Sequence 13, Application US/10734719  
; Publication No. US20060166317A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/734, 719  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US 09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-734-719-13

```
Alignment Scores:
Pred. No.: 2,82e-154 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-734-719-14 (1-291) x US-10-734-719-13 (1-873)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAAACAGTGTTCACACCCCTAAATTCCTCTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTCATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACATATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAAGAAAAATTAACACTAC 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATATCTATCTTCTAGTAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAATAAAAATTAAGAAAAATGTTTATTAAGTTGATAAAGATCTATTAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAGCATTTATTTTCAAGGAAAA 873

RESULT 2
US-10-734-719-8
```

```
; Sequence 8, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10734,719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-734-719-8
```

```
Alignment Scores:
Pred. No.: 9,43e-154 Length: 876
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 6 Gaps: 0

US-10-734-719-14 (1-291) x US-10-734-719-8 (1-876)
```

```
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAAACAGTGTTCACACCCCTAAATTCCTCTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
```

```
Db 481 TCATTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTCGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATACGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAATAAAACCTATATGCTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAAGAAAAATACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACCTACATCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTAATTTCAAGGAAAA 873
```

RESULT 3

```
US-10-734-719-11
; Sequence 11, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/734, 719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-734-719-11
```

```
Alignment Scores:
Pred. No.: 3.98e-153 Length: 873
Score: 1553.00 Matches: 289
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 99.3% Indels: 0
DB: 6 Gaps: 0
```

US-10-734-719-14 (1-291) x US-10-734-719-11 (1-873)

```
Qy 1 MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
```

```
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTGGTCTTCTTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACAATTTAATCCAAAAATCAAGAATATGAGCCGAACATAATATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAACTTTGTAAAAAATTTTACGATTATTTTCTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnIleLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACCAAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTAATCTTTCCGGGAATTTGATTTTATCAAAATCGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAATAAAATTAATTTGCTTATGCTTCTTAACAGTCTTTTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTTAAATTTTATCATACAAAGAAAAATACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACCTACATCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTAATTTCAAGGAAAA 873
```

RESULT 4

```
US-10-734-719-4
; Sequence 4, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/734, 719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
```

; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3(alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-734-719-4

Alignment Scores:  
Pred. No.: 4,32e-150 Length: 876  
Score: 1524.00 Matches: 284  
Percent Similarity: 97.9% Conservative: 1  
Best Local Similarity: 97.6% Mismatches: 6  
Query Match: 97.4% Indels: 0  
DB: 6 Gaps: 0

US-10-734-719-14 (1-291) x US-10-734-719-4 (1-876)

QY 1 MetLysLysVallelleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAAAATGATTGATGATTTAGATGCAATCAATTTTATTGGAAGATAAATACTAT 120  
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTGGGTAATAAATTCACAGATATTATTACAATCCTGGCTCTTTTGTGAACAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAAAATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGCTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCATTTGGATATGATTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTACGAAATTTATCTCAATCAAGAAATTAACCTCAGGAGTCTATATGTCGAGTAGCT 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCTAGGATACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTGGCTCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCACATATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTACAAAAATAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAAAATTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240

DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATACTACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATACTACTACTCTCTAGTAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780  
QY 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAATAAATAAATAAAGAAAAATATTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 5  
US-10-734-719-2  
; Sequence 2, Application US/10734719  
; Publication No. US20060166317A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/734,719  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US 09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3(alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-734-719-2

Alignment Scores:  
Pred. No.: 8,91e-150 Length: 876  
Score: 1521.00 Matches: 282  
Percent Similarity: 98.3% Conservative: 4  
Best Local Similarity: 96.9% Mismatches: 5  
Query Match: 97.3% Indels: 0  
DB: 6 Gaps: 0

US-10-734-719-14 (1-291) x US-10-734-719-2 (1-876)

QY 1 MetLysLysVallelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTGGTAAAAAATGCAAGGAGTATTTTACAATCCTTATCTTTTTTTGAAACAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAATTAATATGTTCTAAT 240

```
Qy 81 TyrAsnGlnAlaHisLeuGluAenGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCT 300
Qy 101 AsnAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAAGCTTTATGCTTATTTAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnAenGlnLeuThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyLysLeuAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCTAAACAAAAATCTTTTAAAAATGGCTCTCTAAATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrLysGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAAATAAACTATATGTCTTATGCTTACAGCTCTTTTGTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGGCGCAATTTAAATTTCAATTTTATCATACAGAAATAAAATACTACTACT 720
Qy 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACCTACTACTCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysLysLysLysLysGluAenValTyrTyrLysLysLysLysLysLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATAAAGAAAAATATTTTATCAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTTCAAGGAAAA 873
```

## RESULT 6

```
US-10-734-719-1
; Sequence 1, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: Gilbert, Michel
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10734, 719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
```

## US-10-734-719-1

```
Alignment Scores:
Pred. No.: 2,36e-148 Length: 11474
Score: 1521.00 Matches: 282
Percent Similarity: 98.3% Conservative: 4
Best Local Similarity: 96.9% Mismatches: 5
Query Match: 97.3% Indels: 0
DB: 6 Gaps: 0
```

## US-10-734-719-14 (1-291) x US-10-734-719-1 (1-11474)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAAGTTATTTATTTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTTATTTCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAAAATGATTTGATGTATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
Qy 41 LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6168 CTTGGTAAAAATGCAAGGAGTATTTTACAACTCTATCTTTTGTGAACAATACTACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTTAAAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTATACGATTATTTTCT 6347
Qy 101 AsnAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGAAATTTTAAATGCTTTATTTTAA 6407
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTTACAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGCTGCGATAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrIleuSerGlyLysPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACCTAAACAAAAATCTTTTAAAAATGGCTCTCTAAATTTTAAAAAT 6587
Qy 181 AspArgSerHisTyrLysGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 6588 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 6647
Qy 201 GluLysThrTyrLysLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTACAAAAATAAACTATATTTGCTTCTTACAGTCTTTTACGAAATTTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGGCGCAATTTTAAATTTCAAAATTTTATCATACAGAAAAATAAATACTACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATATCTACTACTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTTAATTTT 6827
Qy 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAAATAAAAAATAAAGAAAAATATTTTATCAAGTTGATAAAAGATCTATTAAAGATTA 6887
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCATTTATTTTCAAGGAAAA 6920
```

## RESULT 7

```

US-10-734-719-6
; Sequence 6, Application US/10734719
; Publication No. US20060166317A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/734, 719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816, 028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118, 213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495, 406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-734-719-6

Alignment Scores:
Pred. No.: 1.26e-148 Length: 876
Score: 1510.00 Matches: 278
Percent Similarity: 98.3% Conservative: 8
Best Local Similarity: 95.5% Mismatches: 5
Query Match: 96.5% Indels: 0
DB: Gaps: 0

US-10-734-719-14 (1-291) x US-10-734-719-6 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGATGCAATCAATTTTATTGGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTATTTTTACAATCCTAGTCTTTTGTGAAACAATAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAAATCAAGATATGAGACCGAAGTAATCATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAGAACTTTTACGATTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGAAATTCATGCTTATTTTAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTTATTTCATCAATCAAGAAATTAACCTCAGGGGTCTATATGTCACAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATCTTCGGGAATTCATCAAAATGGATCA 480

```

```

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAAATCTTTTAAATTTGGCTCCTCTAAATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuLysCysProAsnSerLysLeuAlaAsnPhe 220
Db 601 GAAAAAATTCACGAAATAAAGCTATATGTATTATGCTCTACAGTCITTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGGCGCAATTTTAAATTTTATCATACAAGAAAAATACTATATACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCATACCTTCTAGTGAGCTTATGGAATAATTACAAAAATATTAAATTTT 780
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATATTATTACAAAGTTGATAAAGATCTATTAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 8
US-11-174-307B-3183
; Sequence 3183, Application US/11174307B
; Publication No. US20060143729A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; TITLE OF INVENTION: USEFUL FOR MODIFYING PLANT CHARACTERISTICS
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174, 307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583, 671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583, 781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583, 651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 3183
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1659)
; OTHER INFORMATION: Ceres cDNA ID no. 21572748
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1659)
; OTHER INFORMATION: Ceres CLONE ID no. 710520
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1659)
; OTHER INFORMATION: Also known as Ceres cDNA ID no. 23378862
US-11-174-307B-3183

Alignment Scores:
Pred. No.: 0.0502 Length: 1659
Score: 112.00 Matches: 72
Percent Similarity: 39.0% Conservative: 62
Best Local Similarity: 20.9% Mismatches: 90
Query Match: 7.2% Indels: 120
DB: Gaps: 22

```

US-10-734-719-14 (1-291) x US-11-174-307B-3183 (1-1659)

```
Qy 9 AsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPhe 28
Db 377 CACTCGCGCGGCTATTCAAGCTAGCTATGAGAAATGGAGAAGAAGTTCAAGGTTTAT 436
Qy 29 -----ArgCysAsnGlnPheGluAspLysTyrTyrLeuGlyLys 43
Db 437 ATATACCCAGATGGGATCCCAACACGCTTCTACGAGCGCGGAGGAGCTCACCGGAAG 496
Qy 44 LysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLys 63
Db 497 -----TATGCCAGTGAAGGCTATTCTTCAGAAAT----- 526
Qy 64 HisLeuIleGlnAsnGlnGluThrGluLeuIleMetCysSerAsnTyrAsnGln 83
Db 527 -----ATCAGAGACAGTCGCTCCGACCCGAA-----AATCCGATGAG 565
Qy 84 AlaHisLeu----- 86
Db 566 GCACACCTCTCTTCATCCCATCTCGTGTGCACAAAGATCGGTGGCAAGGCGACAGTTAT 625
Qy 87 GluAsn-----GluAsnPheValLysThrPheTyrAspTyrPhePro--- 100
Db 626 GAGAATATGACAATAATTGTGCAAAATATATGTGAGAGCTTGATATCCAAATATCCTTAT 685
Qy 101 ---AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 686 TGGAAATAGAACCTTGGGTGCTGATCACTTC-----TTTGTG 721
Qy 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal----- 134
Db 722 ACTTGTCTATGATGT-----GGTGTGAGGCAACAGAGACTTGATTTCTTGTGAAG 775
Qy 135 -----TyrMetCysAlaValAlaIleAlaLeuGlyTyr-----LysGlu 147
Db 776 AATTCCATTCCAGCGGTGTCTCCCGCAGCTATGATGTGATTCATTCCACATAAAGAT 835
Qy 148 IleTyrLeu-----SerGlyIleAspPheTyrGln 157
Db 836 GTGCTCTCCCTCAAGTGTACAGCCATTTGCTCTCTGCTGTGTGGGAATGATATAGAA 895
Qy 158 AsnGlySerSerTyrAlaPhe-----AspThrLys----- 167
Db 896 AACAGGACTACTCTTGATTTCTGGCGTGTGCATCGGAACCTCTAAATCAGAGTTATTCTT 955
Qy 168 -----GlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArg---Ser 183
Db 956 GCACGTGTGTGGGAAATGACACAGACTT-----GATATTCAACAATAGAAATTAGT 1009
Qy 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGlyLysThr 203
Db 1010 AGGCTACTGGGCAT-----CTAGTGTACCAAGAAAGATTT 1045
Qy 204 TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223
Db 1046 TACAGGAGTAGAAG-----TTTTGTATATGTCCTGGTGATCAGAGTTAATAGTCTCGGATA 1102
Qy 224 AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyr----- 239
Db 1103 GCAGACTCTATCCATTATGGTGTATTCTCTGTGATATTGTCAAATYACTATGACCTTCCT 1162
Qy 240 ThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsn 259
Db 1163 TTCAATGATATTCCT-----GACTGGGAACAAATTTGCC----- 1195
Qy 260 PheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArg 279
Db 1196 -----GTTGTACTCAAGGAGAGTGTATTTACCAGCTT-----AAACAAATCCTCAAA 1243
Qy 280 LeuProSerAsp 283
Db 1244 AATATATCAGAT 1255
```

## RESULT 9

US-10-734-719-1/c  
; Sequence 1, Application US/10734719  
; Publication NO. US20060166317A1

## ; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-000111US

; CURRENT APPLICATION NUMBER: US/10734,719

; CURRENT FILING DATE: 2003-12-11

; PRIOR APPLICATION NUMBER: US 09/816,028

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/118,213

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406

; PRIOR FILING DATE: 2000-01-31

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 11474

; TYPE: DNA

; ORGANISM: Campylobacter jejuni

; FEATURE:

; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384

; OTHER INFORMATION: including LOS biosynthesis locus

US-10-734-719-1

## Alignment Scores:

Pred. No.:	0.847	Length:	11474
Score:	110.50	Matches:	76
Percent Similarity:	40.8%	Conservative:	43
Best Local Similarity:	26.0%	Mismatches:	100
Query Match:	7.1%	Indels:	74
DB:	6	Gaps:	16

US-10-734-719-14 (1-291) x US-10-734-719-1 (1-11474)

```
Qy 43 LysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeu 62
Db 6212 AAAAAGATAGGATGTAAATACTGCTTGCATTTTACCAGATAGTATTATCTTC 6153
Qy 63 Lys---HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn--- 80
Db 6152 AAAATAAAATGTATTACATCTAAATA-CATCAAAATCATTTGGTAGCTTGAATAATCAA 6094
Qy 81 -----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPheValLysThrPheTyrAspTy 98
Db 6093 TTTCCTTTAAACTTGGTCCATTTCCAGCAATAAATACTTT-----TTTCATGATTA 6043
Qy 98 rPheProAspAlaHisLeuGlyTyrAspPheLys----- 110
Db 6042 TTTTCCATCAGATTATTTTAAAGATTTTATCAAAATAGATTATTTATGTTATAATAA 5983
Qy 111 -GlnLeuLysGluPheAsnAlaTyr-PhelysPheHisGluIleTyr-----P 126
Db 5982 TTATAATAAAATAAGGAATAAATATGTTTAAAAATTTCAATCATCTTACCACCTTATAATG 5923
Qy 126 heAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrL 146
Db 5922 TGGAAACAATATATAGCAAGGCAATAGAAAGCTGT-----ATCAATCAGACTTTTAA 5872
Qy 146 ysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspT 166
Db 5871 AAGATATAGAAATATTTGTAGTTGAT-----GATT 5842
Qy 166 hrLysGlnGluAsnLeuLysLeuAlaProAspPhe---LysAsnAspArgSerHisT 185
Db 5841 GTGGAAATGATAATAGTATAATATAGCCAAAGAAATACTCTTAAAAAGACAAAGAAATAA 5782
```

```
QY 185 yriledGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrL 205
Db 5781 AAATAATCCAAATGAA-----AAAACCTTAGCTCTTTAAAGACGAAGATAG 5734
QY 205 ys---lleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuA 224
Db 5733 AAGGTGTGAAGTA-----GCAAACTCTCTTATATAATGTTTTTAGATCCTG 5686
QY 224 laProAsnLeuAsnSerAsnPhe-----llelleGlnGluLysAsnA 238
Db 5685 ATGATTATTTTGGAACATAATGCTGTGAAGAGTGTATAAAAAATTTTAGATGAACAGATG 5626
QY 238 sn-----TyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysP 254
Db 5625 AAGTTGATTAGTGTGTTTCAATGCTATTGTTGAAAGTAATGTTATTTCAATATAAAAGT 5566
QY 254 heSerLysAsnIleAsnPhe-----LysLysIleLysIleLysG 267
Db 5565 TTGACTTTAATCTCGGTTTTTATAGCAAAAAAGAGTTTGTAAAAAAAATTTATTGCAAGA 5506
QY 267 luAsnValTyrTyr-----LysLeuIleLysAsp----- 276
Db 5505 AAAATTTATATTGGAATCTATGTGGGGAAACTTATAAGAAAAAGAAATTTATTAGAAGCTT 5446
QY 277 -----LeuLeuArgLeuProSerAspIleLys 285
Db 5445 TTCCGAGTTTAAAGACTCGAAGAAAGATGTTAAA 5414
```

## RESULT 10

```
US-11-216-545-3113
; Sequence 3113, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Soybeans.
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216, 545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3113
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Glycine max
US-11-216-545-3113
```

```
Alignment Scores:
Pred. No.: 0.162 Length: 1614
Score: 107.00 Matches: 72
Percent Similarity: 38.7% Conservative: 61
Best Local Similarity: 20.9% Mismatches: 91
Query Match: 6.8% Indels: 120
DB: Gaps: 22
```

US-10-734-719-14 (1-291) x US-11-216-545-3113 (1-1614)

```
QY 9 AsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPhe 28
Db 352 CACTCGCGCGGTTCAGCTGAACATAGAGAAATGGAAGAAAGTTCAAGGTTTAT 411
QY 29 -----ArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLys 43
Db 412 ATATACCCAGATGGGATCCTAACACGCTTCTACAGACGCCGAGGAGCTCACCGGAAAG 471
QY 44 LysCysLysThrValPheTyrProAsnPhePheGluGlnTyrTyrThrLeuLys 63
Db 472 -----TACGCCAGGAAGGCTATTCTTCAGAAC----- 501
```

```
QY 64 HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGln 83
Db 502 -----ATCAGACAGAGTCGCTTCTGCACCGAA-----AATCCGATGAG 540
QY 84 AlaHisLeu----- 86
Db 541 GCACACCTCTCTTCTTCCATCTCGTGTCAACAGATCGTGGCAAGGGCACGAGTTAT 600
QY 87 GluAsn-----GluAsnPheValLysThrPheTyrAspTyrPhePro--- 100
Db 601 GAGAAATATGACGATAAATTGTACAAAACATGTGGAGAGCTTGATATCCAAATATCCCTTAT 660
QY 101 ---AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 661 TGAACACAGAACCTTGGGTGCTGATCACTTC-----TTTGTGTC 696
QY 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal----- 134
Db 697 ACTTGTGATGATGTT-----GGCGTGAGGGCAACACAGAGGACTTGAGTTTCTTGTGAAG 750
QY 135 -----TyrMetCysAlaValAlaIleAlaLeuGlyTyr-----LysGlu 147
Db 751 AATTCCATTGAGCTGTGTGCTCTCCCGCAGCTATGATGTTGGATTCATTCGCCATAAAGAT 810
QY 148 IleTyrLeu-----SerGlyIleAspPheTyrGln 157
Db 811 GTTGCTCTCCCTCAAGTGTACAGCCATTTGCTCTTCTCGCTGGGGGAATGATATAGAA 870
QY 158 AsnGlySerSerTyrAlaPhe-----AspThrLys----- 167
Db 871 AACAGAACTACTCTTGGATTTTGGCGCGGTCACTGAACCTCTAAAAATCAGAGTTATTCTT 930
QY 168 -----GlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArg---Ser 183
Db 931 GCACGTGTGTGGGAAAAATGACAGAACTT-----GATATTTCAAAACAATAGAAATTAGT 984
QY 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr 203
Db 985 AGGCTACTGGGCAT-----CTAGTGTAACCAAGAGAGATTT 1020
QY 204 TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223
Db 1021 TACAGGAGTAAG--TTTGTATATGTCCTGGTGGATCAGAGGTTAATAGTCTCGGATA 1077
QY 224 AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyr----- 239
Db 1078 GCAGACTCTATCCATTATGGTGTATTCTCTGTGATATTGCTAATTAATCTATGACCTTCCT 1137
QY 240 ThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsn 259
Db 1138 TTCAATGATATTCTT-----GACTGGAAACAAATTTGCT----- 1170
QY 260 PheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArg 279
Db 1171 -----GTTGTACTCAAGGAGAGTGATGATACAGCTT---AAACAATCTCAAA 1218
QY 280 LeuProSerAsp 283
Db 1219 AATATATACAGAT 1230
```

## RESULT 11

```
US-10-471-571A-1901
; Sequence 1901, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471, 571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
```





Db 466 GAGATTTCACCGCTTCTCTG----- 486  
Qy 213 ProAsnSerLeuLeuAlaAsnPhelLeuLeuAlaProAsnLeuAsnSerAsnPhelLe 232  
Db 487 TCTAATAACATGTTGTAACACAGTATTAAAGACTCAGCCTAAATCTGAACAGCCACGCTAAA 546  
Qy 233 lIeGInGluLys--AsnAsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTyr 251  
Db 547 GTGATGGAAGAGTTATCAACTTTCGCCACTTCTGGCTCAATGCCATATCAGCATGCTTAT 606  
Qy 252 GlyLysPheSerLysAsnIle-AsnPhelLysLysIleLysLysGluAsnValTyr 270  
Db 607 GAAAGTGTAAGATTATCAATTACAAGCTCCTGTGTTTGAACAAATCCAGTACAT 664

## RESULT 13

US-10-471-571A-251  
; Sequence 251, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927WO  
; CURRENT APPLICATION NUMBER: US/10/471,571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 251  
; LENGTH: 1953  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-251

Alignment Scores:  
Pred. No.: 0.609 Length: 1953  
Score: 102.50 Matches: 52  
Percent Similarity: 36.0% Conservative: 28  
Best Local Similarity: 23.4% Mismatches: 66  
Query Match: 6.6% Indels: 77  
DB: Gaps: 10

## US-10-734-719-14 (1-291) x US-10-471-571A-251 (1-1953)

Qy 14 LysGluIleAspTyrSerArgLeuProAsnAspPhe-----AspValPheArg----- 29  
Db 338 AAACAGATTGGTTATCAAAACATGCA-AATGATTACGATAGTGATACATATAAAACACAT 396  
Qy 30 -----CysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGly 42  
Db 397 ATGTCCTCGTTATGAAPCTTTATCGAATCAATTTTGAACAAATTAGAAGCATATCAATATGAA 456  
Qy 43 LysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyr----- 59  
Db 457 AGTAAATAATTAACACAGTACTTTACGGGTAAATTTTAGTGAAGAGATTTCATATAACCT 516  
Qy 60 -----TyrThrLeuLysHisLeuIleGlnAsn 68  
Db 517 ATCAATGATTTAGCGGTGGCCAAAAACACAGTTTATCTTAGCTCAATGCTATTAAAC 576  
Qy 69 GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsn 88  
Db 577 -----GAACCTGATTATTACTTTTAGATGAA----- 603  
Qy 89 GluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhe 108  
Db 604 -----CCTACTAACCACTTAGATTAGAAACG 630  
Qy 109 PheLysGlnLeuLysGluPheAsnAlaTyrPheLys-----PheHis 122  
Db 631 ACAAGTGCTTGAAGATTATCTACGTTATTTTAAAGGTGCAATCGTCATCATCAGCCAT 690

Qy 123 GluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAla 142  
Db 691 GATCGTTACTTTTAGATAAAATAGTTACTCAAATTTAT-----GATGTGCT 738  
Qy 143 LeuGlyTyrLysGluIleTyrLeuSerGlyIle---AspPheTyrGlnAsnGlySerSer 161  
Db 739 TTAGGTGATGCAACCGCTATGTTGGTAATTACGAGGAATTTATACAGCAACGGGATTTA 798  
Qy 162 TyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp 181  
Db 799 TATTATCAAAAACGAATGCAAGATAT----- 825  
Qy 182 ArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGlu---PheLeu 200  
Db 826 -----GAAAGTCAACAACGAGAAATAAAACGATTAGAACTTTTCTGT 867

## RESULT 14

US-11-217-529-5949  
; Sequence 5949, Application US/11217529  
; Publication No. US20060099612A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNTORY LIMITED  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHISA  
; APPLICANT: KODAMA, YUKIKO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIKO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: S-38-285  
; CURRENT APPLICATION NUMBER: US/11/217,529  
; CURRENT FILING DATE: 2005-09-02  
; PRIOR APPLICATION NUMBER: US 10/932,182  
; PRIOR FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 5949  
; LENGTH: 1248  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-11-217-529-5949

Alignment Scores:  
Pred. No.: 0.389 Length: 1248  
Score: 102.00 Matches: 70  
Percent Similarity: 34.7% Conservative: 50  
Best Local Similarity: 20.2% Mismatches: 108  
Query Match: 6.5% Indels: 118  
DB: Gaps: 17

## US-10-734-719-14 (1-291) x US-11-217-529-5949 (1-1248)

Qy 23 AsnAspPheAspVal-----PheArgCysAsnGlnPheTyrPheGluAspLysTyr 39  
Db 22 AACGAGCCTGATGTTGGATTTCCTCATCTCAAGTTGAATCAGCAACTATTTCATATCCTTAC 81  
Qy 40 TyrLeuGlyLysLysCys-----LysThrValPheTyrThrProAsnPhe 54  
Db 82 GAATCTATCTCAATATCGTTAAACATACTCAACCACTATATATAAGAACCAAAATTT 141  
Qy 55 PhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu 74  
Db 142 TTACACGAACATATCATCAGTACGCTGAACGAAATCTTTGAGCACAATGATGTCGAGCATGAT 201  
Qy 75 LeuIleMetCysSerAsnTyrAsnGln----- 83  
Db 202 AGATTAGCACTCGCAAAAAATAAACGAACTGATCAGAAAAAATTGACCATATATGAAAAAGTTC 261  
Qy 84 -----AlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAsp 97

Db 262 TTGAATATTCAAGCTAAATCACATTGGCAAAATCTAAATAGAAATCAAGAAA----- 312  
Qy 98 TyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPhe----- 115  
Db 313 -----AGATTGGAGTTTTTCCATGAACCTGAAAGATATCAAAATCC 351  
Qy 116 -----AsnAlaTyrPheLys 120  
Db 352 CGAAACATGAGATTCACCTTAATGAAATAACCCAGATATGAGAAATAAATGATTCGG 411  
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCys----- 137  
Db 412 TGGTATCAAAAGTTACAAAC-----ATTTGATTGGCGACTACTTAACTAGGAATAAC 465  
Qy 138 AlaValAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGln 157  
Db 466 CCGATCAATATAATTCAGAGACAAATCAACACTGGAATTCAGGGTAGTGTCTTGGAAA 525  
Qy 158 AsnGlySerTyrAlaPheAspThrLysGlnGluAsnLeuLeu----- 172  
Db 526 -----CAAAACCAATGGAGGATCTCATCGATTATGATGTCTC 564  
Qy 173 -----LysLeuAlaProAspPheLysAsnAspArgSer----- 183  
Db 565 TTGGAGGCCAATAAGATTCCACATCATTTAAACGACACGATCTCTTACCCTTAATT 624  
Qy 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr 203  
Db 625 TCATGGGTAAATGAAATAAAAGACTTTGACCAAAACATCTTCTATGCTAGAAATTCCAA 684  
Qy 204 TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu-----AlaAsnPheIle 221  
Db 685 GCAAGACTACAAAGATACATT-----GAACTATTGAAAGCTGGCAACTATATC 732  
Qy 222 Glu-----LeuAlaProAsnLeuAsnSerAsnPhe----- 231  
Db 733 GATGCTATCTTTGTTCCAGAAATACCTTTTATGAGAGATTAATTTCAAGAT 792  
Qy 232 -----IleIleGlnLysAsnAsnTyrThrLysAspIle 243  
Db 793 TTAAATATTAGCATCTGGACTTCTGATATTTATCAAGTATTGCAACAGCAAAA----- 846  
Qy 244 LeuIleProSerSerGluAlaTyrGlyLysPhe-----SerLysAsn 257  
Db 847 -----CCAAATTCATCTCTGGGTCTAGACTTGATCTGAAGATATCAAAATCAAAAAC 900  
Qy 258 IleAsnPheLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277  
Db 901 CTACAATGAATAAGACCAATATTTCAACATTTCTTCCATAAATGT----- 948  
Qy 278 LeuArgLeuProSerAsp 283  
Db 949 -----CTACCGAGGAT 960

## RESULT 15

US-11-217-529-6121  
; Sequence 6121, Application US/11217529  
; Publication No. US20060099612A1  
; GENERAL INFORMATION:  
; APPLICANT: SUNTORY LIMITED  
; APPLICANT: NAKAO, YOSHIHIRO  
; APPLICANT: NAKAMURA, NORIHISA  
; APPLICANT: KODAMA, YUKIKO  
; APPLICANT: FUJIMURA, TOMOKO  
; APPLICANT: ASHIKARI, TOSHIHIKO  
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS  
; FILE REFERENCE: S-38-285  
; CURRENT APPLICATION NUMBER: US/11/217,529  
; CURRENT FILING DATE: 2005-09-02  
; PRIOR APPLICATION NUMBER: US 10/932,182  
; PRIOR FILING DATE: 2004-09-02  
; NUMBER OF SEQ ID NOS: 197023  
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 6121  
; LENGTH: 2175  
; TYPE: DNA  
; ORGANISM: Saccharomyces pastorianus  
US-11-217-529-6121

Alignment Scores:  
Pred. No.: 0.789 Length: 2175  
Score: 102.00 Matches: 71  
Percent Similarity: 36.0% Conservatives: 43  
Best Local Similarity: 22.4% Mismatches: 115  
Query Match: 6.5% Indels: 88  
DB: 8 Gaps: 14  
US-10-734-719-14 (1-291) x US-11-217-529-6121 (1-2175)  
Qy 6 IleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArgLeuProAsnAspPhe 25  
Db 802 CTGGGTTCAAAGGGGCCCAATTTTAAAGAAGTC-----ATTCCACTTGAA--- 846  
Qy 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCys 45  
Db 847 -----AACGAAGACCTTAAATAATTAATGAAAAAATGT 879  
Qy 46 -----LysThrValPheTyrThrProAsnPhePhePheGluGlnTyr 59  
Db 880 TTTCGATTGTCAGTTTTTAAAGATGTGTACTAGTACGGTTTCTGGACGATCATAACTTT 939  
Qy 60 TyrThrLeuLysHisLeuIleGlnAsnGln----- 70  
Db 940 AATTTAATCTCGAAATCGTAATGATTTAGAAACATGCATCATCGACTTTCTTCAAGTG 999  
Qy 71 -----TyrGluThrGluLeuIleMetCysSerAsn 80  
Db 1000 GGGACTTTCTTAGATAGGCTGATAGAGCTTTATGATGACTGATCATTCACGATAGCTCC 1059  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 1060 -----GTAGAGAGAAAAGTTCTGTAACAAAACGTAAGACCGTATCAGA 1104  
Qy 101 -----AspAlaHisLeuGlyTyrAspPhe 108  
Db 1105 TTACTCCAACAATGTGTTCAAATGCCATTAAATTTAGACGCGCATGACCGTTCCAAAGTTT 1164  
Qy 109 PheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGln 128  
Db 1165 TATAAGACACTTGTCCGA---AAGGGTCTATTCAAAGTCTTAGATTATGCAITTCACATG 1221  
Qy 129 ArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148  
Db 1222 GAAACTGATAGTAACGTTTCGCTACCGATACCTATCATTCATCCATCATCGAGCAT 1281  
Qy 149 TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGln 168  
Db 1282 GACATATTGCTGATTCAACCGTTCAAATGAAAACCTCTTTC-----AAACAA 1329  
Qy 169 GluAsnLeuLysLeuAlaProAspPheLys-----AsnAspArgSerHisTyrIleGly 187  
Db 1330 CAACAC-----AAATTGGCCCGGATGACAAATGTTCTTACACGAAGCAATTCACAGGAT 1383  
Qy 188 HisSerLysAsnThrAspIleLysAlaLeuGluPheLeu----- 200  
Db 1384 TACAACCTCAGCAGCTGACTCCAAGGTATTTAATAATCTTCAACGATTTCTTTAAACCGAC 1443  
Qy 201 -----GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsn 214  
Db 1444 AAAAGTCCCGGTTACGGCAACAGTTGTGCAAGCGCTAAACACCGTTCTTCCACCAGAA 1503  
Qy 215 SerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGln 234  
Db 1504 GGATGCTCTCAGTAAT-----GGTGAAGGTTCTGATATGATATCATGAGT 1545  
Qy 235 GluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPhe 254

```
Db      1546 AGATCCCAACACGAATCAAAAAT-----ACCTCTGAAGACTTCCCAAAATTC 1593
Qy      255 SerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyr 271
Db      1594 GGTATCGATTGAATTCGGATTCAATTAAATTG---GACAATTACAACACTAC 1641
```

Search completed: August 21, 2006, 16:05:21  
Job time : 205.5 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: August 21, 2006, 15:05:59 ; Search time 6599.5 Seconds  
(without alignments)  
3866.527 Million cell updates/sec  
Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIAGNPSLKEIDYSR.....KLIKDLRLPSDIKHVFKGK 291

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 86534536 seqs, 29229259966 residues  
Total number of hits satisfying chosen parameters: 173069072  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/abss/ABSSWEB.spool/US10734719/runat.21082006.141821.12465/app\_query.fasta.1  
-DB=Pending Patents.NA.Main -OPMT=fastap -SUFFIX=P2N.rnpnp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0  
-USER=US10734719 @CGN 1.1 10839 @runat.21082006.141821.12465 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6 -FGAEXT=7  
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXI=7

Database :

Pending Patents.NA.Main:

1:	/EMC_Celerra_SIDS3/ptodata/2/pna/PCTUSA_COMB.seq.*
2:	/EMC_Celerra_SIDS3/ptodata/2/pna/PCTUSB_COMB.seq.*
3:	/EMC_Celerra_SIDS3/ptodata/2/pna/PCTUSC_COMB.seq.*
4:	/EMC_Celerra_SIDS3/ptodata/2/pna/US075_COMB.seq.*
5:	/EMC_Celerra_SIDS3/ptodata/2/pna/US076_COMB.seq.*
6:	/EMC_Celerra_SIDS3/ptodata/2/pna/US077_COMB.seq.*
7:	/EMC_Celerra_SIDS3/ptodata/2/pna/US078_COMB.seq.*
8:	/EMC_Celerra_SIDS3/ptodata/2/pna/US079_COMB.seq.*
9:	/EMC_Celerra_SIDS3/ptodata/2/pna/US080_COMB.seq.*
10:	/EMC_Celerra_SIDS3/ptodata/2/pna/US081_COMB.seq.*
11:	/EMC_Celerra_SIDS3/ptodata/2/pna/US082_COMB.seq.*
12:	/EMC_Celerra_SIDS3/ptodata/2/pna/US083_COMB.seq.*
13:	/EMC_Celerra_SIDS3/ptodata/2/pna/US084_COMB.seq.*
14:	/EMC_Celerra_SIDS3/ptodata/2/pna/US085_COMB.seq.*
15:	/EMC_Celerra_SIDS3/ptodata/2/pna/US086_COMB.seq.*
16:	/EMC_Celerra_SIDS3/ptodata/2/pna/US087_COMB.seq.*
17:	/EMC_Celerra_SIDS3/ptodata/2/pna/US088_COMB.seq.*
18:	/EMC_Celerra_SIDS3/ptodata/2/pna/US089_COMB.seq.*
19:	/EMC_Celerra_SIDS3/ptodata/2/pna/US090_COMB.seq.*
20:	/EMC_Celerra_SIDS3/ptodata/2/pna/US091_COMB.seq.*
21:	/EMC_Celerra_SIDS3/ptodata/2/pna/US092_COMB.seq.*
22:	/EMC_Celerra_SIDS3/ptodata/2/pna/US093_COMB.seq.*
23:	/EMC_Celerra_SIDS3/ptodata/2/pna/US094_COMB.seq.*
24:	/EMC_Celerra_SIDS3/ptodata/2/pna/US095A_COMB.seq.*
25:	/EMC_Celerra_SIDS3/ptodata/2/pna/US095B_COMB.seq.*
26:	/EMC_Celerra_SIDS3/ptodata/2/pna/US095C_COMB.seq.*
27:	/EMC_Celerra_SIDS3/ptodata/2/pna/US096A_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	100.0	873	47	US-10-303-161-13 Sequence 13, Appl
2	1564	100.0	873	60	US-10-734-719-13 Sequence 13, Appl
3	1564	100.0	873	61	US-10-820-536-13 Sequence 13, Appl
4	1564	100.0	873	61	US-10-821-573-13 Sequence 13, Appl

```
5 1564 100.0 873 61 US-10-821-604-13 Sequence 13, Appl
6 1564 100.0 873 61 US-10-830-825-13 Sequence 13, Appl
7 1564 100.0 873 61 US-10-830-997-13 Sequence 13, Appl
8 1564 100.0 873 61 US-10-845-408-13 Sequence 13, Appl
9 1564 100.0 873 61 US-10-845-412-13 Sequence 13, Appl
10 1564 100.0 873 61 US-10-846-219-13 Sequence 13, Appl
11 1564 100.0 873 61 US-10-847-983-13 Sequence 13, Appl
12 1564 100.0 873 61 US-10-850-125-13 Sequence 13, Appl
13 1564 100.0 873 61 US-10-850-807-13 Sequence 13, Appl
14 1564 100.0 873 63 US-10-961-882-13 Sequence 13, Appl
15 1564 100.0 873 63 US-10-962-235-13 Sequence 13, Appl
16 1564 100.0 873 63 US-10-962-334-13 Sequence 13, Appl
17 1559 99.7 876 67 US-10-303-161-8 Sequence 8, Appl
18 1559 99.7 876 60 US-10-734-719-8 Sequence 8, Appl
19 1559 99.7 876 61 US-10-820-536-8 Sequence 8, Appl
20 1559 99.7 876 61 US-10-821-573-8 Sequence 8, Appl
21 1559 99.7 876 61 US-10-821-604-8 Sequence 8, Appl
22 1559 99.7 876 61 US-10-830-825-8 Sequence 8, Appl
23 1559 99.7 876 61 US-10-830-997-8 Sequence 8, Appl
24 1559 99.7 876 61 US-10-845-408-8 Sequence 8, Appl
25 1559 99.7 876 61 US-10-845-412-8 Sequence 8, Appl
26 1559 99.7 876 61 US-10-846-219-8 Sequence 8, Appl
27 1559 99.7 876 61 US-10-847-983-8 Sequence 8, Appl
28 1559 99.7 876 61 US-10-850-125-8 Sequence 8, Appl
29 1559 99.7 876 61 US-10-850-807-8 Sequence 8, Appl
30 1559 99.7 876 63 US-10-961-882-8 Sequence 8, Appl
31 1559 99.7 876 63 US-10-962-235-8 Sequence 8, Appl
32 1559 99.7 876 63 US-10-962-334-8 Sequence 8, Appl
33 1553 99.3 873 47 US-10-303-161-11 Sequence 11, Appl
34 1553 99.3 873 60 US-10-734-719-11 Sequence 11, Appl
35 1553 99.3 873 61 US-10-820-536-11 Sequence 11, Appl
36 1553 99.3 873 61 US-10-821-573-11 Sequence 11, Appl
37 1553 99.3 873 61 US-10-821-604-11 Sequence 11, Appl
38 1553 99.3 873 61 US-10-830-825-11 Sequence 11, Appl
39 1553 99.3 873 61 US-10-830-997-11 Sequence 11, Appl
40 1553 99.3 873 61 US-10-845-408-11 Sequence 11, Appl
41 1553 99.3 873 61 US-10-845-412-11 Sequence 11, Appl
42 1553 99.3 873 61 US-10-846-219-11 Sequence 11, Appl
43 1553 99.3 873 61 US-10-847-983-11 Sequence 11, Appl
44 1553 99.3 873 61 US-10-850-125-11 Sequence 11, Appl
45 1553 99.3 873 61 US-10-850-807-11 Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-10-303-161-13
; Sequence 13, Application US/10303161
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II

; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-161-13
Alignment Scores: 2.19e-235 Length: 873
Pred. No.: 1564.00 Matches: 291
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 47

US-10-734-719-14 (1-291) x US-10-303-161-13 (1-873)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGTATTATTAGATGTAATCAATTTTATTATTGAAGATAACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTCTTTGACCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACAATATATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAATAACTTTTACGATTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACCAAAATTTATTTTCAATCAAAAGAAATACCTTCAGGGGCTATATATGTGCGAGTACC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACATATATCGGACATAGTAAAAATACATATATAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAACATATATTTGCTTCTTAATAGTCTTTTAGCAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAspTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAAATAACTACTACT 720
Qy 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAAGATATACTCATACCTCTTAGGAGCTTATGGAATAATTTTCAAAAAATTTTAAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAAATTAAGAAATTTGTTTATTACAAAGTTGATAAAGATCTATTAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTTCAAGGAAAA 873
```

RESULT 2  
US-10-734-719-13  
; Sequence 13, Application US/10734719  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/734, 719  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR FILING DATE: US 09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-734-719-13

Alignment Scores:  
Pred. No.: 2,19e-235 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 60 Gaps: 0

US-10-734-719-14 (1-291) x US-10-734-719-13 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
DB 61 CTACCAATGATTTCATGATGATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTTGGTAAAAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTAAATCCAAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATCAGAAAATGAAAATTTGTAAAACTTTTACGATTATTTTCCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCATTTGGGATGATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnAcqIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTACAGAAATTTATTTCATCAAAAGAAATTAACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACCAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTCAATTTTATCATACAGAAAAAATAACTACACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATTAAGATCTATTAGATTTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTATATAAGCATTAATTTCAAGAGAAAA 873

RESULT 3  
US-10-820-536-13  
; Sequence 13, Application US/10820536  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-820-536-13

Alignment Scores:  
Pred. No.: 2,19e-235 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-820-536-13 (1-873)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60

```
Db      1  ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAAGG 60
QY      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAATGATTTGATGATTAATAGATTAATCAATTTATTTTGAAGATAAATACTAT 120
QY      41  LeuGlyLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      121  CTGGTAAATAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
QY      61  ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db      181  ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGCACTAATTAATGTTCTAAT 240
QY      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTATTTCCT 300
QY      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATATGATTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY      121  PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY      141  IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
QY      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAsnPheLysAsn 180
Db      481  TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATGCTTCACACTATATCGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
QY      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTTACAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTT 660
QY      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
```

## RESULT 4

```
US-10-821-573-13
; Sequence 13, Application US/10821573
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
```

```
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CscII) from C. jejuni O:36
US-10-821-573-13
```

## Alignment Scores:

Pred. No.:	2,19e-235	Length:	873
Score:	1584.00	Matches:	291
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	61	Gaps:	0

US-10-734-719-14 (1-291) x US-10-821-573-13 (1-873)

```
QY      1  MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1  ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAAGG 60
QY      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAATGATTTGATGATTAATAGATTAATCAATTTTATTTGAGATAAATACTAT 120
QY      41  LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      121  CTGGTAAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACTAC 180
QY      61  ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db      181  ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGCACTAATTAATGTTCTAAT 240
QY      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTATTTCCT 300
QY      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATATGATTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY      121  PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY      141  IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
QY      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAsnPheLysAsn 180
Db      481  TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATGCTTCACACTATATCGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
QY      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTTACAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTT 660
QY      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
```



```
Db 661 ATAGAACTAGCGCAAAATTAAATTTCAAAATTTTATCATCAAGAAAAATAACTACACT 720
Qy 241 LysAspIleuLeuProSerSerGluAlaTyGlyLysPheSerLysAsnIleAsnPh 260
Db 721 AAAGATATACTCATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyTyTyLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAGATT 840
Qy 281 ProSerAspIleLysHisTyTyTyPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTATTTTCAAGGAAAA 873

RESULT 5
US-10-821-604-13
; Sequence 13, Application US/10821604
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,604
; CURRENT FILING DATE: 2004-04-08
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-821-604-13

Alignment Scores:
Pred. No.: 2,19e-235 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-821-604-13 (1-873)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyTySerArg 20
Db 1 ATGAAAAAGTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyTyPheGluAspLysTyTy 40
Db 61 CTACCAATGATTTTGCATGTTAGATGTAATCAATTTTATTTGAAGATAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyTyThrProAsnPhePheGluGlnTyTyTy 60
Db 121 CTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyTyGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAAATCAAGAAATATAGACCGCAACTAATATATGTGTTCTAAT 240
```

```
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyTyAspTyTyPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTCTAAAAAATCTTTTACGATTAATTTCTCT 300
Qy 101 AsnAlaHisLeuGlyTyTyAspPhePheLysGlnLeuLysGluPheAsnAlaTyTyPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAAAACAACCTTAAAGAAATTTAAATGCTTTATTTTAAA 360
Qy 121 PheHisGluIleTyTyPheAsnGlnArgIleThrSerGlyValTyTyMetCysAlaValAla 140
Db 361 TTTACGAAAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATATGTGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyTyLysGluIleTyTySerGlyIleAspPheTyTyGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTC 480
Qy 161 SerTyTyAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTATGATACCAAAACAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyTyIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyTyLysIleLysLeuTyTyCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyTyThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATCAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuLeuProSerSerGluAlaTyTyGlyLysPheSerLysAsnIleAsnPh 260
Db 721 AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyTyTyTyLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTTAAAGAAAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAGATT 840
Qy 281 ProSerAspIleLysHisTyTyPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTATTTTCAAGGAAAA 873

RESULT 6
US-10-830-825-13
; Sequence 13, Application US/10830825
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/830,825
; CURRENT FILING DATE: 2004-04-24
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
```

; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-830-825-13

## Alignment Scores:

Pred. No.: 2,196-235 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-830-825-13 (1-873)

Qy	1	MetLysLysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAATGATTTTGATGATTAGATGTAATCAATTTTATTGAGATAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	121	CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAAAACATTTAATCCAAAATCAAGATATGAGACCGAACTAAATATGTTCTTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAGAACTTTTACGATTATTTTCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGluLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA	360
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTATTTCATCAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAATCTTACAAATATAAATCTATATGCTTATGCTTAATAGTCTTTTACGAAATTTT	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCAATTTTAAATTTCAATTTTATCATACAAGAAAAATACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATACCTACCTCTTAGTGAGCTTATGGAATTTTTCAAAAAATATTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAATAAATAAATAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATT	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGATATAAGCATTTATTTCAAGGAAAA	873

## RESULT 7

US-10-830-997-13

; Sequence 13 Application US/10830997  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/830,997  
; CURRENT FILING DATE: 2004-04-22  
; PRIOR APPLICATION NUMBER: US/10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-830-997-13

## Alignment Scores:

Pred. No.: 2,196-235 Length: 873  
Score: 1564.00 Matches: 291  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-830-997-13 (1-873)

Qy	1	MetLysLysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAATGATTTTGATGATTATGATGTAATCAATTTTATTGAGATAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	121	CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA	360
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTATTTCATCAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160

Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
 Db 481 TCCTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
 Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
 Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
 Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 Db 601 GAAAAAATCTTACAAAATAAAATATATGCTTATGTCTCTATAGTCTCTTTAGCAATTTT 660  
 Qy 221 IleGluLeuAlaProAsnLeuAsnSerAspPheIleGlnGluLysAsnAsnTyrThr 240  
 Db 661 ATAGACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATTTTAACTACACT 720  
 Qy 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 Db 721 AAAGATATATCATACCTCTAGTGAGCTTATGGAATTTTCAAAAAATTTAAATTT 780  
 Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeuArgLeu 280  
 Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
 Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 8

US-10-845-408-13  
 ; Sequence 13, Application US/10845408  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/845,408  
 ; CURRENT FILING DATE: 2004-05-12  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn ver. 2.1  
 ; SEQ ID NO 13  
 ; LENGTH: 873  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(873)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-845-408-13

Alignment Scores:  
 Pred. No.: 2,19e-235 Length: 873  
 Score: 1564.00 Matches: 291  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-845-408-13 (1-873)

Qy 1 MetLysLysValIleIleAlaGlyProSerLysGluIleAspTyrSerArg 20  
 |||

Db 1 ATGAAAAAAGTATTATTCTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60  
 Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
 Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
 Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60  
 Db 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTCGAGCAATACTAC 180  
 Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
 Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACATAATTTATGTGTTCTAAT 240  
 Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300  
 Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATCTTATTAAAA 360  
 Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 Db 361 TTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
 Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
 Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
 Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
 Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
 Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
 Db 601 GAAAAAATCTTACAAAATAAAATATATGCTTATGTCTCTTAAATAGTCTTTTACCAATTTT 660  
 Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATTTTAACTACACT 720  
 Qy 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
 Db 721 AAAGATATATCATACCTCTAGTGAGCTTATGGAATTTTCAAAAAATTTAAATTTT 780  
 Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeuArgLeu 280  
 Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
 Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
 Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 9

US-10-845-412-13  
 ; Sequence 13, Application US/10845412  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/845,412  
 ; CURRENT FILING DATE: 2004-05-12  
 ; PRIOR APPLICATION NUMBER: US/10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ;

```
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-412-13

Alignment Scores:
Pred. No.: 2,19e-235 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-845-412-13 (1-873)
QY 1 MetLysLysVallelleAlaGlyAsnGlyProSerLeuLysGluileAspTyrSerArg 20
DB 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGTGATTTAGATGTAATCAATTTTATTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuileGlnAsnGlnGluTyrGluThrGluLeuileMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCAACTAATTAATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY 121 PheHisGluileTyrPheAsnGlnAsGilleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACACAAATTTATTTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluileTyrLeuSerGlyileAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProaspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrileGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysileLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheileGlnGluLysAsnAsnTyrThr 240
```

```
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTATCATCAAGAAAAAATAACTACTACT 720
QY 241 LysAspLysLeuileProSerSerGluAlaTyrGlyLysPheSerLysAsnLeuPhe 260
DB 721 AAAGATATACTACTACTCTTCTAGTGGCTTATGAAAAATTTTCAAAAAATATTAAATTT 780
QY 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLysLeuileLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAAAATTTAAAGAAATGTTTATTACAGTTGATAAAAGATCTATTAGATT 840
QY 281 ProSerAspLysLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 10
US-10-846-219-13
; Sequence 13, Application US/10846219
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; PRIOR FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-846-219-13

Alignment Scores:
Pred. No.: 2,19e-235 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-846-219-13 (1-873)
QY 1 MetLysLysVallelleAlaGlyAsnGlyProSerLeuLysGluileAspTyrSerArg 20
DB 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGTGATTTAGATGTAATCAATTTTATTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACTAC 180
QY 61 ThrLeuLysHisLeuileGlnAsnGlnGluTyrGluThrGluLeuileMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCAACTAATTAATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
QY 121 PheHisGluileTyrPheAsnGlnAsGilleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACACAAATTTATTTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluileTyrLeuSerGlyileAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProaspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrileGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysileLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheileGlnGluLysAsnAsnTyrThr 240
```

```
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATATTTCTCT 300
Qy 101 AsplahisLeuGlyTyrAspPhePheGlyGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnAlaTyrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Qy 201 GlulysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTATAGTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleLeuGlnLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCCAATTTAAATTTCAAAATTTTATCATACAAGAAAAATTAACACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTCTAGGAGCTTATGGAATAATTTTCAAAAAATTTATTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
```

## RESULT 11

```
US-10-847-983-13
; Sequence 13, Application US/10847983
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-847-983-13
```

```
Alignment Scores:
Pred. No.: 2,19e-235 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 61 Gaps: 0
US-10-734-719-14 (1-291) x US-10-847-983-13 (1-873)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTTATGCTGGAATCGAACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGTATGATGTTATGATGTTATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTGCGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATTTATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTTATTTCT 300
Qy 101 AsplahisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTTAAATGCTTATTTTAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnAlaTyrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAAACCTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Qy 201 GlulysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleLeuGlnLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCCAATTTAAATTTCAAAATTTTATCATACAAGAAAAATTAACACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTCTAGGAGCTTATGGAATAATTTTCAAAAAATTTATTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
RESULT 12
US-10-850-125-13
```

```

; Sequence 13, Application US/10850125
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,6-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-850-125-13

Alignment Scores:
Pred. No.: 2,19e-235 Length: 873
Score: 1564.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 61 Gaps: 0

US-10-734-719-14 (1-291) x US-10-850-125-13 (1-873)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCGTGAAATGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTGATGTATTAGATGATCAATTTATTTTGAAGATTAATACTAT 120

Qy 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAAAACATTTAATCAAAATCAAGAATATGACCGAAGTAATATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCAATTTGGGATGATGATTTTAAAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTATTTCAATCAAAAGAAATTAATCTTTCCGGGAATTTGATTTTATCAAAATGGGCA 480

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180

```



Qy	261	LysLysLysLeIeLysLysGluAsnValTyrTyrLysLeulleLysAspLeuleuArgLeu	280
Db	781	AAAAAAAAATAAAAYTTAAAGAAAAATGTATTATCAAGTTTCATAAAGATCTATTAAAGATTA	840
Qy	281	ProSerAspLleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTAGATATAAAGCATTTATTTCAAAGGAAA	873

```

RESULT 15
US-10-962-235-13
; Sequence 13, Application US/10962235
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/962,235
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-962-235-13

```

Alignment Scores:	2,19e-235	Length:	873
Pred. No.:	Score:	Matches:	291
	1564.00	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	100.0%	Gaps:	0
DB:	63		

US-10-734-719-14 (1-291) x US-10-962-235-13 (1-873)

Qy	1	MetLysLysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAATAAAAGTTATTATTCTGTGGAAATGACCAAGTTTTAAAAGAAATTGATTATTCAAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAGATGATTTTGATGTATTAGATGTATACTAATTTTATTTTGAAGATAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePhePheGluGlnTyrTyr	60
Db	121	CTTGGTAAAAAATGCMAAACAGTGTFTTTACACCCCTTAATTCCTCTTGAGCAATACTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAACAATTTAATCCAAAATCAGAATAATGAGCCGAACCTAATTATGTGTTCTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluLysGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCCAGCTCATCTAGAAAAATGAAAAATTTTGTAANAACCTTTTACGATTATTTTCCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTTTTTTAAAACACTTAAGAATTTAATGCTTATTTTAA	360



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 15:30:50 ; Search time 124 Seconds

(without alignments)  
3175.070 Million cell updates/sec

Title: US-10-734-719-14

Perfect score: 1564

Sequence: 1 MKKVIAGNSGLKIDYSR.....KLIKDLRLPSDIKHYFKGK 291

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1281442 seqs, 450983614 residues

Total number of hits satisfying chosen parameters: 2562884

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/abss/ABSSWEB spool/US10734719/runat 21082006 141825 12524/app query.fasta\_1  
-DB=Pending Patents NA New -OFMT=fasta -SUFFIX=P2N.rnpn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -HOST=abss08  
-USER=US10734719 @CGN 1 1.163 @runat 21082006 141825 12524 -NCPU=6 -ICPU=3  
-NO WMAP -NEG SCORES=0 -WAIT -DSPLOG=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:

1: /EMC\_Celerra\_SIDS3/ptodata/2/pna/PCT NEW COMB seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US06 NEW COMB seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US07 NEW COMB seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US08 NEW COMB seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US09 NEW COMB seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US10 NEW COMB seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US11 NEW COMB seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US60 NEW COMB seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	123	7.9	6128	US-10-433-793A-69	Sequence 69, Appl
C 2	104.5	6.7	83391	US-10-433-793A-123	Sequence 123, Appl
C 3	100	6.4	960	US-11-434-137-4645	Sequence 4645, Ap
C 4	100	6.4	960	US-11-434-137-4645	Sequence 4645, Ap
C 5	100	6.4	960	US-11-434-137-4645	Sequence 4645, Ap
C 6	100	6.4	960	US-11-434-137-4645	Sequence 4645, Ap
C 7	100	6.4	960	US-11-434-137-4645	Sequence 4645, Ap
C 8	100	6.4	34688	US-10-433-793A-90	Sequence 90, Appl
C 9	100	6.4	2160266	US-11-434-137-10967	Sequence 10967, A

10	100	6.4	2160266	7	US-11-434-184-10967	Sequence 10967, A
11	100	6.4	2160266	7	US-11-434-203-10967	Sequence 10967, A
12	100	6.4	2160266	7	US-11-434-127-10967	Sequence 10967, A
13	100	6.4	2160266	7	US-11-434-199-10967	Sequence 10967, A
14	98.5	6.3	5984	6	US-10-433-793A-23	Sequence 23, Appl
15	98	6.3	2814	1	PCT-US06-15100-300	Sequence 300, App
16	98	6.3	5984	6	US-10-433-793A-24	Sequence 24, Appl
17	96	6.1	7341	6	US-10-433-793A-139	Sequence 139, App
18	96	6.1	9095	6	US-10-433-793A-91	Sequence 91, Appl
19	96	6.1	81440	7	US-11-158-863-659	Sequence 659, App
20	95	6.1	3510	7	US-11-027-891A-1748	Sequence 1748, Ap
21	95	6.1	3510	7	US-11-027-879A-1748	Sequence 1748, Ap
22	94.5	6.0	6456	6	US-10-582-705-399	Sequence 399, App
23	93.5	6.0	849	1	PCT-US06-08555A-357	Sequence 357, App
24	93.5	6.0	849	1	PCT-US04-07095B-357	Sequence 357, App
25	93	5.9	2856	6	US-10-551-492-4	Sequence 4, Appli
26	93	5.9	20579	6	US-10-433-793A-103	Sequence 103, App
27	92.5	5.9	5611	6	US-10-433-793A-99	Sequence 99, Appl
28	91.5	5.9	43445	6	US-10-712-795A-334	Sequence 334, App
29	91	5.8	4001	6	US-10-582-705-278	Sequence 278, App
30	91	5.8	6070	6	US-10-433-793A-159	Sequence 159, App
31	91	5.8	14121	6	US-10-712-795A-3	Sequence 3, Appli
32	90.5	5.8	4491	6	US-10-582-705-330	Sequence 330, App
33	90.5	5.8	13870	7	US-11-431-708-6087	Sequence 6087, A
34	90.5	5.8	13870	7	US-11-475-062-10831	Sequence 10831, A
35	90.5	5.8	13870	8	US-60-812-072-32	Sequence 32, Appl
36	90.5	5.8	13960	7	US-11-431-708-6085	Sequence 6085, Ap
37	90.5	5.8	13960	7	US-11-431-708-6086	Sequence 6086, Ap
38	90.5	5.8	13960	7	US-11-475-062-10829	Sequence 10829, A
39	90.5	5.8	13960	7	US-11-475-062-10830	Sequence 10830, A
40	90.5	5.8	13960	8	US-60-812-072-30	Sequence 30, Appl
41	90.5	5.8	13960	8	US-60-812-072-31	Sequence 31, Appl
42	90.5	5.8	13993	6	US-10-712-795A-318	Sequence 318, App
43	90.5	5.8	440684	1	PCT-US03-41389-114	Sequence 114, App
44	90	5.8	945	7	US-11-027-891A-639	Sequence 639, App
45	90	5.8	945	7	US-11-027-879A-639	Sequence 639, App

#### ALIGNMENTS

RESULT 1  
US-10-433-793A-69/c  
; Sequence 69 Application US/10433793A  
; GENERAL INFORMATION:  
; APPLICANT: Schacht, Oliver  
; TITLE OF INVENTION: DIAGNOSIS OF DISEASES ASSOCIATED WITH ANGIOGENESIS  
; FILE REFERENCES: 47675-185  
; CURRENT APPLICATION NUMBER: US/10/433,793A  
; CURRENT FILING DATE: 2003-06-06  
; PRIOR APPLICATION NUMBER: PCT/EP01/14320  
; PRIOR FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: DE 10061338.1  
; PRIOR FILING DATE: 2000-12-06  
; NUMBER OF SEQ ID NOS: 212  
; SEQ ID NO 69  
; LENGTH: 6128  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (3742)  
; OTHER INFORMATION: unknown base  
US-10-433-793A-69

Alignment Scores:  
Pred. No.: 0.0146  
Score: 123.00  
Percent Similarity: 33.6%  
Best Local Similarity: 23.6%  
Query Match: 7.9%  
DB: 6  
Length: 6128  
Matches: 92  
Conservative: 39  
Mismatch: 108  
Indels: 154  
Gaps: 18

```
US-10-734-719-14 (1-291) x US-10-433-793A-69 (1-6128)
QY 31 AsnGlnPheTyrPheGluAspLysTyr-TyrLeuGlyLysLysCysLysThrValPheTy 50
Db 2197 AATCGCTTCCTTTTCAAAACAAATAACGTAATTTTAACTAAATAA---AAATCCTACACACA 2141
QY 50 rThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnAsnGlnGl 70
Db 2140 CACAAACACTATATAATAAATAAATAATAC---AACTTAATTC-AATCAACA 2091
QY 70 uTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAs 90
Db 2090 A---AAATTCATATAACAATATAAACAACATCAATCACTACCTCGAAAATATCAA 2034
QY 90 nPhe-----ValLysThrPheTyrAs 97
Db 2033 CATTCCTACTATTAAATAATAATTTTAAATAATTAATAAACATCACTCAAAACCATTTATAT 1974
QY 97 pTyr----- 98
Db 1973 CTACTATTATAACATACCAACAAATAAATAATCATCTACATACACAAACAAAT 1914
QY 99 ----PhePro-----AspAlaHisLe 104
Db 1913 AAATTTCCAAACATACCAACCAACCAACCAATCAATCATATCGTATCTCTTAATAT 1854
QY 104 uGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluL 124
Db 1853 TTAACCTTCATTTTACAATTTATTAATAAATAAATTAATTAATTTTCTTTTACCACACTTT 1794
QY 124 eTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGl 144
Db 1793 AAATCAATATCAATATATTTCGAAATAATCTATAC----- 1760
QY 144 yTyrLysGluIleTyrLeuSer-----GlyIleAspPheTyrGlnAsnGlySerSery 162
Db 1700 CTTCACTATATTCTAAACTTAAATTTTAAACACCGAATCACTC-ACGCTATATAT 1642
QY 175 aProAspPhe---LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIl 194
Db 1641 CCCATCAGCTTTTAAACACCGAA-----ATAAACGAATCAACGAATAAATAAATAAAT 1594
QY 194 eLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCys----- 210
Db 1593 TAAACCAATC-----CTCACTAACAGATATAAACCAATCTCTACTA-AAATATAA 1541
QY 210 ----- 210
Db 1540 AATTATCCGAACGTATATCGAACGCTATATATCCCACTACTCGAAAACTAAACAAAA 1481
QY 211 -----LeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaPro----- 225
Db 1480 AAATACGTTAAACCCAAATAACGAAACTTTACATATAAACCGAATTTCCAGCCACTACACTC 1421
QY 226 -AsnLeuAsnSerAsnPheIleIleGln-GluLysAsnAsnTyrThrLysAspIleLeuI 245
Db 1420 CAACTTAAACGACGAAAAATAAACTCAATCTCAAAAAAATAAATAAATAAATAAATAAATTC 1361
QY 245 leProSerSerGluAlaTyrGlyLysPheSerLysAsnIle----- 258
Db 1360 TAAATCAATAATAAATAAATAAATAAATTTCAAAAAATACCATAATATAAATAAATAAATA 1301
QY 259 -----AsnPheLysLysIleL 264
Db 1300 CAAAACTAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1241
QY 264 ysIle-----LysGluAsn----- 268
Db 1240 AATTCGCCACCTCCAAACAAAAATAAATAATATATATTCCTACCCCTTTAAATTCCT 1181
QY 269 --ValTyrTyrLys-----LeuIleLysAspLeuLeuArgLeup 281
Db 1180 AAACATATTTTAAACCTTTAAACCTTATAATATTTCTTATATATTCCTTATATAATATTC 1121
QY 281 roSerAspIleLysHisTyrPheLys 289
Db 1120 CATCATTTCTTAAACATATTTTCAAAA 1095
RESULT 2
US-10-433-793A-123/c
; Sequence 123, Application US/10433793A
; GENERAL INFORMATION:
; APPLICANT: Schacht, Oliver
; TITLE OF INVENTION: DIAGNOSIS OF DISEASES ASSOCIATED WITH ANGIOGENESIS
; FILE REFERENCE: 47675-185
; CURRENT APPLICATION NUMBER: US/10/433, 793A
; CURRENT FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: PCT/EP01/14320
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: DE 10061338.1
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 123
; LENGTH: 83391
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4324..4325, 4327..4328, 4330..4331, 4333..4334, 4336..4337)
; OTHER INFORMATION: unknown base
; NAME/KEY: unsure
; LOCATION: (4339..)
; OTHER INFORMATION: unknown base
US-10-433-793A-123
Alignment Scores:
Pred. No.: 24.5 Length: 83391
Score: 104.50 Matches: 65
Percent Similarity: 37.4% Conservative: 51
Best Local Similarity: 21.0% Mismatches: 111
Query Match: 6.7% Indels: 83
DB: Gaps: 16
US-10-734-719-14 (1-291) x US-10-433-793A-123 (1-83391)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 7268 ATCAAAAAAATATTTTATAAACTAAACATAACAAATTTTCACTACCCCTAATATTCACCT 7209
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 7208 ATATATCATCTACTACGACTTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAA 7149
QY 41 LeuGlyLysLysCysLysThrValPheTyrThrProAsnPhePheGluGlnTyr 60
Db 7148 ATCAATACAAA-----AAATTTACA 7128
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 7127 AATATACCTCCTTCAACAACTAAATATATCTAAATATCTAAATATCTAAATATCTAAAT 7080
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 7079 ---CAAAAAATAAATTTACAAAAATAATTCATTTTCACTCCCACTTTAAATAATTAATTTAC 7023
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 7022 CATCTACATTTAAATAACAACTAAATTT----- 6996
```

```
QY 121 PheHisGluIleTyrPhe-----AsnGlnArgIleThrSerGlyValTyrMetCys 137
||||| : : : : : ||| : : :
Db 6995 TTCACAAATTATTTACATAAACATATAACACATAAAACT-----6951
QY 138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLysSerGlyIleAspPheTyrGln 157
||||| : : : : : ||| : : :
Db 6950 -----ACATATAAAAAATATACAAAAAACTAACAAATCAATATTTA 6909
QY 158 AsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAsp 177
||||| : : : : : ||| : : :
Db 6908 AACACACACT-----ACAAAAAAATAAAAAAT-----AAAAATTTCCCTTAAT 6870
QY 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu 197
||||| : : : : : ||| : : :
Db 6869 TTTAAAAAT-----ATTCAATTAACATAAAAACTAAAAAACTTCT--ATAATTTA 6819
QY 198 GluPheLeuGluIlysthrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu 217
||||| : : : : : ||| : : :
Db 6818 CAAATTAATAACAACTTTATACAAAAAATT-----ATAATTTA 6777
QY 218 AlaAsnPheIleGluLeuAlaProAsn-----LeuAsnSerAsnPheIle 232
||||| : : : : : ||| : : :
Db 6776 AAATTTTAAAAATTTCTCCTAACTCTCTTTAAATTTCTCACATTCAAATTAATCTC 6717
QY 233 IleGlnGluLysAsnAsnTyrThrLys-----241
Db 6716 -----AAAAAAACATCAAAAAAATCCATTTTAAATCAATAATAATATTTTA 6666
QY 242 AspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLys---AsnIleAsnPhe 260
||||| : : : : : ||| : : :
Db 6665 AACATACTTTACATCCCAATAAATCAATTTAAACACCTATACACATAAACATACACATT 6606
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeu---IleLysAspLeuArg 279
||||| : : : : : ||| : : :
Db 6605 AAAAACTTAAATAACTACAAAAAAACAAATACCAACAAATATATCTCAATAATCAAT 6546
QY 280 LeuProSerAspIleLysHisTyrPheLys 289
||||| : : : : : ||| : : :
Db 6545 TTTTCCATCA-----TACTACCTAAAA 6525
```

## RESULT 3

```
US-11-434-137-4645
; Sequence 4645, Application US/11434137
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Massignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Herve
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,137
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 4645
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-11-434-137-4645
```

Alignment Scores:

```
Pred. No.: 0.359 Length: 960
Score: 100.00 Matches: 75
Percent Similarity: 38.9% Conservative: 60
Best Local Similarity: 21.6% Mismatches: 102
Query Match: 6.4% Indels: 110
DB: 7 Gaps: 22

US-10-734-719-14 (1-291) x US-11-434-137-4645 (1-960)

QY 1 MetLysLysValIleIleAlaGly-----AsnGlyProSer 12
: : : : : ||| : : :
Db 7 ATTAATAAGCTATTGTTGCAGGCTTCTGCTTTTATTATTACCTAGCGAGTCAGCA 66
QY 13 LeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGln 32
: : : : : ||| : : :
Db 67 TATGCTCATCAGATGTCAGAAAGTT-----ATTGATGAA 102
QY 33 PheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysThrValPheTyrThrPro 52
: : : : : ||| : : :
Db 103 ACTTATGTTCAACCTGATTGCTTAGGT-----132
QY 53 AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72
: : : : : ||| : : :
Db 133 -----TATTCGCTA-----AATCAAGAACAGCGT 156
QY 73 ThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHis-----Leu 86
: : : : : ||| : : :
Db 157 GCACAGACATTACAACTTCTGAATTAATGATGAATCTAGACACACGAAAGTTAAACGCTA 216
QY 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
: : : : : ||| : : :
Db 217 AATACTAGCTCATATGCTAAA--ATAATGAATATAGCTGATGATGCAAGCATCCAATTA 273
QY 107 AspPhePheLysGlnLeuLysGluPhe-----115
: : : : : ||| : : :
Db 274 TATTTCATCTGTTAAATAAAAAAATTAGGGTCTTAACGATACCTTAGCAGTAAACATTGT 333
QY 116 -----AsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSer 132
: : : : : ||| : : :
Db 334 ACACCTGAGAATATTCTAAAGTTACAGAAGATGATGTCGCAATGTCGAGTACATTA 393
QY 133 GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys-----GluIleTyr 149
: : : : : ||| : : :
Db 394 GGTATTGAGCATGCAACGATATCAGTGCCTCCAATAAAAGTAACATGGGAGAGTGCT 453
QY 150 LeuSerGlyIleAspPheTyr-----GlnAsnGlySerSerTyrAlaPheAspThr 166
: : : : : ||| : : :
Db 454 CTTGAGGTATC--TATTATTTCATAGAGAAAAATGGAGCAAGTGTATCATCAGAAAAAT 510
QY 167 Lys-----GlnGluAsnLeuLysLeuAlaPro---AspPheLysAsn---Asp 181
: : : : : ||| : : :
Db 511 AAACAACCTAGCACAGAAGAGTTAAGTACTTTATCTGTTATTAATGCTGAGAAATAGGGT 570
QY 182 ArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu--- 197
: : : : : ||| : : :
Db 571 AAGGAAGTTATGACCGCGACACAGCTTAATGTAGCTTTAACCGATATTAAAGTCAGCCGTA 630
QY 198 -----GluPheLeuGluLysThrTyr 204
: : : : : ||| : : :
Db 631 GCTAAAGGAGGATCTGATCTTTCAAAAGATGATATTTCGAAAAAATGTTGAGGAAACA--- 687
QY 205 LysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu--- 223
: : : : : ||| : : :
Db 688 ---TTGAAAAATTTATCATCTT---GATAACCTGTACCGGAAAAATCAATAATTTAAAT 741
QY 224 -----AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsn----- 237
: : : : : ||| : : :
Db 742 GTTAATTTTGCAGTTAATCTTTTCAAAAGTAATGTCAATCAAAAAATAGTATTTCAGCAAC 801
QY 238 -----AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255
: : : : : ||| : : :
Db 802 ACCTTGATTAATTTGAAAGACACACATTGTT-----TCGAAAGCAGGATCTAAATTTAAA 855
```

```
QY 256 Lys---AsnIleAsnPhelYsLysIleLysIleLysGlu----- 267
Db 856 AATATCAATGTCATTTTAATGCTATATAAAGCTGTGTAATCAGGTAAAGGATTTTAGCT 915

QY 268 AsnValTyrTyrLysLeuLeu 274
Db 916 AATATTGGCAACAATTTGTT 936

RESULT 4
US-11-434-184-4645
; Sequence 4645, Application US/11434184
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,184
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 4645
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-11-434-184-4645

Alignment Scores:
Pred. No.: 0.359 Length: 960
Score: 100.00 Matches: 75
Percent Similarity: 38.9% Conservative: 60
Best Local Similarity: 21.6% Mismatches: 102
Query Match: 6.4% Indels: 110
DB: 7 Gaps: 22

US-10-734-719-14 (1-291) x US-11-434-184-4645 (1-960)

QY 1 MetLysValIleIleIleAlaGly-----AsnGlyProSer 12
Db 7 ATTAATAAGCTATGTTTTCAGGCCTTTCCTTTTATTATTATCCCTAGCGAGTCCAGCA 66

QY 13 LeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGln 32
Db 67 TATGCTGCATCAGATGTTCAAGAGT-----ATTGATGAA 102

QY 33 PheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysThrValPheTyrPro 52
Db 103 ACTATGTCCTCAACCTGATGCTTAGGT----- 132

QY 53 AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72
Db 133 -----TATTCGCTA-----AATCAAGAACACGCGT 156

QY 73 ThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHis-----Leu 86
Db 157 GCACAGACATTACAACCTTCTGAATTTATGTAATCTAGACACGAAAGTTAAACCGCTA 216

QY 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 217 AATACTAGCTCATATGCTAAA---ATAATGAATATAGCTGATGATGCAAGCATCCAATTA 273
```

```
QY 107 AspPhePhelYsGlnLeuLysGluPhe----- 115
Db 274 TATTCATCTCTTAAAAATAAAAAATTAGGGGCTAACGATACCTTAGCAGTAAACATTGTT 333

QY 116 -----AsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSer 132
Db 334 ACACCTGAGAATATTACTAAAGTTTACAGAAGATATGATCGCAATGCTGCAGTTACATTA 393

QY 133 GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys-----GluIleTyr 149
Db 394 GGTTATTGAGCATGCAACGATATCAGTGGCTCTCCAATAAAAGTAACATGGGGAGAGTCT 453

QY 150 LeuSerGlyIleAspPheTyr-----GlnAsnGlySerSerTyrAlaPheAspThr 166
Db 454 CTTCAGGTATC---TATTATTATTAGAGAAAATGGACGAGTATCATCAGAAAAT 510

QY 167 Lys-----GlnGluAsnLeuLysLeuAlaPro-----AspPheLysAsn---Asp 181
Db 511 AAACAACATAGCACAAAGAAGATTAAAGTACTTTATCTGTTATTAATGCTCAGAAATAGGGT 570

QY 182 ArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu--- 197
Db 571 AAGGAAGTTTATGACGCCGACAGCTTAATGCTTTTAAACGGATATTAAAGTCACCGGTA 630

QY 198 -----GluPheLeuGluLysThrTyr 204
Db 631 GCTAAAGGAGGATCTGATCTTTCAAAAGATGATATTCGAAAAATTGTTGAGGAACA--- 687

QY 205 LysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu--- 223
Db 688 ---TTGAAAAAATTATCATCTT---GATAACCGCTGTACCGAAAAATCAAAATAAATTAAAT 741

QY 224 -----AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsn----- 237
Db 742 GTTAATTTTCAGTTAATCTTTACAAAGTAATGTCATCAAAAATAGTATGATTTTCACGAA 801

QY 238 -----AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255
Db 802 ACCTTGAAATAATTTGAAAAGCAACATTTGTT-----TCGAAAGCAGGATCTAAATTTAA 855

QY 256 Lys---AsnIleAsnPhelYsLysIleLysLysGlu----- 267
Db 856 AATATCAATGTCATTTTAATGCTTAATAAAGCTGTGTAATCAGGTAAAGGATTTTAGCT 915

QY 268 AsnValTyrTyrLysLeuLeu 274
Db 916 AATATTGGCAACAATTTGTT 936

RESULT 5
US-11-434-203-4645
; Sequence 4645, Application US/11434203
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,203
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
```

```

; SEQ ID NO 4645
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-11-434-203-4645

Alignment Scores:
Pred. No.: 0.359 Length: 960
Score: 100.00 Matches: 75
Percent Similarity: 38.9% Conservative: 60
Best Local Similarity: 21.6% Mismatches: 102
Query Match: 6.4% Indels: 110
DB: 7 Gaps: 22

US-10-734-719-14 (1-291) x US-11-434-203-4645 (1-960)

QY 1 MetLysLysValIleAlaGly-----AsnGlyProSer 12
DB 7 ATTAAGAAGCTATTGTTGACGGCTGCTTTATTTATCCCTAGGAGTCCAGCA 66
QY 13 LeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGln 32
DB 67 TATGCTGCATCAGATGTTTCAAGAAAGTT-----ATTGATGAA 102
QY 33 PheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysThrValPheTyrThrPro 52
DB 103 ACTTATGTTCAACCTGATTTATGTTAGGT----- 132
QY 53 AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72
DB 133 -----TATTCGCTA-----AATCAAGAACACGGT 156
QY 73 ThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHis-----Leu 86
DB 157 GCACAGACATTACACTTCTGAATATGATGAATCTAGACACAGCAAGATTAAACGCTA 216
QY 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
DB 217 AATACTAGCTCATATGCTAAA---ATAATGAATATAGCTGATGATCAAGCATCCAATTA 273
QY 107 AspPhePheLysGlnLeuLysGluPhe-----115
DB 274 TATTTCATCTGTTAAATAAATAAATAATAGGCTCTAACGATACCTTAGCAGTAAACATTGTT 333
QY 116 -----AsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSer 132
DB 334 ACACCTGAGATATTACTAAAGTTACAGAAGATATGTCGCAATGCTGCGATGTACATTA 393
QY 133 GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys-----GluIleTyr 149
DB 394 GGTATTGACGATCAACGATATCAGTGGCTGCTCCATAAAGATAAAGTGGGAGAGTGCT 453
QY 150 LeuSerGlyIleAspPheTyr-----GlnAsnGlySerSerTyrAlaPheAspThr 166
DB 454 CTTGCAAGTATC---TATTATTCATTAGAGAAAAATGGACAAAGTGTATCATCAGAAAT 510
QY 167 Lys-----GlnGluAsnLeuLysLeuAlaPro---AspPheLysAsn---Asp 181
DB 511 AAACAACACTACACAAGAAGAGTAAAGTACTTTATCTGGTATTAATGCTGAGAATAAGGT 570
QY 182 ArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu--- 197
DB 571 AAGGAAGTTATGACCGCGACAAGCTTAATGTAGCTTTAACGGTATTAATGATCAGCGCTA 630
QY 198 -----GlnPheLeuGluLysThrTyr 204
DB 631 GCTAAAGGAGGATCTGATCTTTCAAAAGATGATATTCGAAAAATTTGTTGAGGAACA-- 687
QY 205 LysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu--- 223
DB 688 ---TTGAAAAATATATCATCTT---GATAACGCTGTACCGAAAAATCAAAATTTAATT 741
QY 224 -----AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsn----- 237

```

```
Db 157 GCACAGACATTACAACCTCTCGAATTATGATGAATCTAGAGACACGAAAGTTAAACGCTA 216
QY 87 GluAsnGluAsnPhelValysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 217 AATACTAGCTCATATGCTAAA---ATAATCAATATATAGCTGATGATCAAGCATCCAATTA 273
QY 107 AspPhePheLeuGlnLeuLysGluPhe----- 115
Db 274 TATTCATCTGTTAAATAATAAAATAGGGTCTAACGATACCTTAGCAGTAACATTGTT 333
QY 116 -----AsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSer 132
Db 334 ACACCTGAGAAATATTACTAAAGTTACAGAAAGATATGTCGCAATGCTGCAGTTACATTA 393
QY 133 GlyValTyMetCysAlaValAlaLeuGlyTyrLys-----GluLeuTyr 149
Db 394 GGTATTGAGCATGCAACGATATCATGGCTCTCAATAAAAGTAACATGGGAGAGTGCT 453
QY 150 LeuSerGlyIleAspPheTyr-----GlnAsnGlySerSerTyrAlaPheAspThr 166
Db 454 CTTCGAGGTATC---TATTATTCAATTAGAGAAAATGGAGCAAGTGTATCATCAGAAAT 510
QY 167 Lys-----GlnGluAsnLeuLeuLysLeuAlaPro-----AspPheLysAsn---Asp 181
Db 511 AAACAACATAGCACACAAGAGAGTTAAGTACTTTATCTGTTATTAATTTACCTAGCGAGTCCAGCA 66
QY 182 ArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu--- 197
Db 571 AAGGAAGGTTATGACGCCGACCAACATTGTTAAGTCTTAACGGATATTAAGTCAGCGCTA 630
QY 198 -----GluPheLeuGluLysThrTyr 204
Db 631 GCTAAAGGAGGATCTGATCTTTCAAAAGATGATATTCGAAAATTTGTTAGGAAACA--- 687
QY 205 LysIleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhelLeuGluLeu--- 223
Db 688 ---TTGAAAAATATCATCTT---GATAACGCTGTACCGGAAATCAATATAATTTAATT 741
QY 224 -----AlaProAsnLeuAsnSerAsnPhelIleGlnGluLysAsn----- 237
Db 742 GTTAATTTTGCAGTTAATCTTTACAAAAGTAATGTCATCAAAAATAGTGATTTACGGAAC 801
QY 238 -----AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255
Db 802 ACCTTGAATAATTGGAAGACACACATTGTT-----TCGAAGCAGGATCTAAATTTAAA 855
QY 256 Lys---AsnIleAsnPhelLysLysIleLysIleLysGlu----- 267
Db 856 AATATCAATGTCATTTTAATGCTAATAAAGCTGTTGAATCAGGTAAGGATTTTAGCT 915
QY 268 AsnValTyTyrIysLeuLe 274
Db 916 AATATTGGCAACAATAATTGTT 936

RESULT 7
US-11-434-199-4645
; Sequence 4645, Application US/11434199
; GENERAL INFORMATION:
; APPLICANT: Telford, John
; APPLICANT: Masignani, Vega
; APPLICANT: Ros, Immaculada Margarit Y
; APPLICANT: Fraser, Claire
; APPLICANT: Tettelin, Hervé
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/11/434,199
; CURRENT FILING DATE: 2006-05-16
; PRIOR APPLICATION NUMBER: US 10/415,182
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: PCT/GB01/04789
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: GB-0026333.5
```

```
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: GB-0028727.6
; PRIOR FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: GB-0105640.7
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 12025
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 4645
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
; US-11-434-199-4645

Alignment Scores:
Pred. No.: 0.359 Length: 960
Score: 100.00 Matches: 75
Percent Similarity: 38.9% Conservatives: 60
Best Local Similarity: 21.6% Mismatches: 102
Query Match: 6.4% Indels: 110
DB: 7 Gaps: 22

US-10-734-719-14 (1-291) x US-11-434-199-4645 (1-960)

QY 1 MetLysLysValIleLeAlaGly-----AsnGlyProSer 12
Db 7 ATTAATAAGCTATTGTTTGCAGGCTTGTCTTTTATTATTATTTACCTAGCGAGTCCAGCA 66
QY 13 LeuLysGluLeuAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGln 32
Db 67 TATGCTGCATCATGATGTTTCAGAAAGTT-----ATTGATGAA 102
QY 33 PheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysThrValPheTyrThrPro 52
Db 103 ACTTATGTTCAACCTGATTAATGCTTAGGT----- 132
QY 53 AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72
Db 133 -----TATTCGCTA-----AATCAAGAACAGCGT 156
QY 73 ThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHis-----Leu 86
Db 157 GCACAGACATTACAACCTCTCGAATTTATGATGAATCTAGAGACACGAAAGTTAAACGCTA 216
QY 87 GluAsnGluAsnPhelValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 217 AATACTAGCTCATATGCTAAA---ATAATGAATATAGCTGATGATGCAAGCATCCAATTA 273
QY 107 AspPhePheLysGlnLeuLysGluPhe----- 115
Db 274 TATTCATCTCTTAATAATAAAATAGGGTCTAACGATACCTTAGCAGTAACATTGTT 333
QY 116 -----AsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSer 132
Db 334 ACACCTGAGAAATATTACTAAAGTTACAGAAAGATATGTCGCAATGCTGCAGTTACATTA 393
QY 133 GlyValTyMetCysAlaValAlaLeuGlyTyrLys-----GluIleTyr 149
Db 394 GGTATTGAGCATGCAACGATATCATGGCTCTCAATAAAAGTAACATGGGAGAGTGCT 453
QY 150 LeuSerGlyIleAspPheTyr-----GlnAsnGlySerSerTyrAlaPheAspThr 166
Db 454 CTTCGAGGTATC---TATTATTCAATTAGAGAAAATGGAGCAAGTGTATCATCAGAAAT 510
QY 167 Lys-----GlnGluAsnLeuLeuLysLeuAlaPro-----AspPheLysAsn---Asp 181
Db 511 AAACAACATAGCACACAAGAGAGTTAAGTACTTTATCTGTTATTAATTTACCTAGCGAGTCCAGCA 570
QY 182 ArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu--- 197
Db 571 AAGGAAGGTTATGACGCCGACCAACATTGTTAAGTCTTAACGGATATTAAGTACCGCTA 630
QY 198 -----GluPheLeuGluLysThrTyr 204
```



Query Match:	6.4%	Indels:	110
DB:	7	Gaps:	22
US-10-734-719-14 (1-291) x US-11-434-137-10967 (1-2160266)			
QY	1	MetLysLysVallelleAlaGly-----AsnGlyProSer	12
DB	941654	ATTAAAAAGCTATTGTTGAGCCCTGCTTTATTTTACCCTAGCGAGTCCAGCA	941713
QY	13	LeuLysGluIleAspTyrSerArgLeuProAsnAspPheValPheArgCysAsnGln	32
DB	941714	TATGCTGCATCAGATGTTCAAGAAAGT-----ATTGATGAA	941749
QY	33	PheTyrPheGluAspLysTyrLeuGlyLysLysCysLysThrValPheTyrThrPro	52
DB	941750	ACTTATGTTCAACCTGATTGCTTAGGT-----	941779
QY	53	AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLysGlnAsnGlnGluTyrGlu	72
DB	941780	-----TATTCGCTA-----AATCAAGAACAGCGT	941803
QY	73	ThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHis-----Leu	86
DB	941804	GCACAGACATTACAACCTTCTGAATTTATGATGAATCTAGACACAGAAAGTTAAACCGCTA	941863
QY	87	GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr	106
DB	941864	AATACTAGCTCATATGCTGCTAAA-----ATAATGAATATAGCTGATGATCAATTA	941920
QY	107	AspPhePheLysGlnLeuLysGluPhe-----	115
DB	941921	TATTCATCTGTTAAATAATAAAAAAATTAGGCTTAACGATACCTTAGCGAGTAAACATTGTT	941980
Alignment Scores:			
Pred. No.:	3.3e+03	Length:	2160266
Score:	100.00	Matches:	75
Percent Similarity:	38.9%	Conservative:	60
Best Local Similarity:	21.6%	Mismatches:	102
Query Match:	6.4%	Indels:	110
DB:	7	Gaps:	22
US-10-734-719-14 (1-291) x US-11-434-184-10967 (1-2160266)			
QY	1	MetLysLysVallelleAlaGly-----AsnGlyProSer	12
DB	941654	ATTAAAAAGCTATTGTTGAGCCCTGCTTTATTTTACCCTAGCGAGTCCAGCA	941713
QY	13	LeuLysGluIleAspTyrSerArgLeuProAsnAspPheValPheArgCysAsnGln	32
DB	941714	TATGCTGCATCAGATGTTCAAGAAAGT-----ATTGATGAA	941749
QY	33	PheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysThrValPheTyrThrPro	52
DB	941750	ACTTATGTTCAACCTGATTGCTTAGGT-----	941779
QY	53	AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLysGlnAsnGlnGluTyrGlu	72
DB	941780	-----TATTCGCTA-----AATCAAGAACAGCGT	941803
QY	73	ThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHis-----Leu	86
DB	941804	GCACAGACATTACAACCTTCTGAATTTATGATGAATCTAGACACAGAAAGTTAAACCGCTA	941863
QY	87	GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr	106
DB	941864	AATACTAGCTCATATGCTGCTAAA-----ATAATGAATATAGCTGATGATCAATTA	941920
QY	107	AspPhePheLysGlnLeuLysGluPhe-----	115
DB	941921	TATTCATCTGTTAAATAATAAAAAAATTAGGCTTAACGATACCTTAGCGAGTAAACATTGTT	941980
US-11-434-184-10967			
; Sequence 10967, Application US/11434184			
; GENERAL INFORMATION:			
; APPLICANT: Telford, John			
; APPLICANT: Masignani, Vega			
; APPLICANT: Ros, Immaculada Margarit Y			
; APPLICANT: Fraser, Claire			
; APPLICANT: Tettelin, Herve			
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM STREPTOCOCCUS GROUPS A & B			
; FILE REFERENCE:			
; CURRENT APPLICATION NUMBER: US/11/434,184			
; CURRENT FILING DATE: 2006-05-16			
; PRIOR APPLICATION NUMBER: US 10/415,182			
; PRIOR FILING DATE: 2003-04-28			
; PRIOR APPLICATION NUMBER: PCT/GB01/04789			
; PRIOR FILING DATE: 2001-10-29			
; PRIOR APPLICATION NUMBER: GB-0026333.5			
; PRIOR FILING DATE: 2000-10-27			
; PRIOR APPLICATION NUMBER: GB-0028727.6			
; PRIOR FILING DATE: 2000-11-24			
; PRIOR APPLICATION NUMBER: GB-0105640.7			
; PRIOR FILING DATE: 2001-03-07			
; NUMBER OF SEQ ID NOS: 12025			
; SOFTWARE: SeqWin99, version 1.02			
; SEQ ID NO 10967			
; LENGTH: 2160266			
; TYPE: DNA			
; ORGANISM: Streptococcus agalactiae			
US-11-434-184-10967			







[illegible]

```
QY 36 GluAspLysTyr-----TyrLeuGlyLysLysCysLysThrValPheTyrThr 51
Db 4527 TATACTATTATAAAATCCAAATCACTTTTCTCTCTCCACTTCTTACCAAACT 4468
QY 52 ProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnAsnGlnGluTyr 71
Db 4467 TTCAATTACTCTTTTACAAATATTTCACTCTCTC-CACGNTATCTATAAC----- 4418
QY 72 GluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPhe 91
Db 4417 -----ACATCAAACTATACAACTACTCAAAACAAACATCACTACTTA 4376
QY 92 ValLys-----ThrPheTyrAspTyr-----PheProAspAla-HisLeu 104
Db 4375 TTAATATATTCGTAATAAACAATATTTCCATTTTATTCAAATTTTCTTTTAAATCAAA 4316
QY 104 uGlyTyrAspPhePheLysGlnLeu----- 112
Db 4315 ATATCTCCAATTTCTAAACAAATTTAATCTTTTAAACAAATAAACAACAACTAAC 4256
QY 113 -----LysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnAr 129
Db 4255 GTTACAAAAAATAATTATAACATATATTCAAAATAACCTTAAAAAATAAATAAATAA 4196
QY 129 gileThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluLeu 149
Db 4195 ACACATCA-----AAACACATTTTACGAAATAAATCTAACTA 4157
QY 149 rLeu-----SerGlyLeuAspPheTyrGlnAsnGlySerSerTyrAla----- 164
Db 4156 CTTAAAAAATAAATAATTACTCTAATTAATCAAAATAAATCAATACACATAATTAAT 4097
QY 164 eAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAsp----- 177
Db 4096 TAATACTCAAAATAAATAATTACTTAAATAATTCTTACAAAAATAATTAACCAATTA 4037
QY 178 -PheLysAsnAsp-----ArgSerHisTyrIleGlyHisSerLysAsnThrAs 193
Db 4036 TTTTAAAAATAAATAATTACTTAATTAATAATATCACTATTTTAACTTCCAAAAATA 3977
QY 193 pIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysPr 213
Db 3976 AATTAATCTNCTCATTAATAATAAC-----ATACTCAA 3941
QY 213 oAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231
Db 3940 TCAAAATCTTACTAACTTACTAATCACTCCCTNAAAAATAATCTTACCAACACTTTT 3881
QY 232 -----IleIleGlnGluLysAsnAsnTyrThrLysAspIleLe 244
Db 3880 AAATTACTTTATAAAATCAACACTAATAATTAACCTTAAACCTTAAACCTTAAAC 3822
QY 244 uIleProSerSerGluAlaTyrGlyLys-----PheSerLysAsn----- 257
Db 3821 ATTAAAAATAAACAATAAATAATCTTAAAAACACTTCTCCAAAAACGATAACCAATA 3762
QY 258 -----IleAsnPheLysLysIleLysIleLysGluAsnValTyrTy 271
Db 3761 ACATAAAAAATACTCGACATTAACATATATAAATAAATAAATAAATAAATAAATAA 3702
QY 271 rLysLeuIle 274
Db 3701 CTACTTCATA 3692
```

## RESULT 15

```
PCT-US06-15100-300
; Sequence 300, Application PC/TUS0615100
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation et al.
; TITLE OF INVENTION: ESCHERICHIA COLI O157:H7 PROTEINS AND USES THEREOF
; FILE REFERENCE: 00786/458W02
; CURRENT APPLICATION NUMBER: PCT/US06/15100
```

```
; CURRENT FILING DATE: 2006-04-20
; PRIOR APPLICATION NUMBER: 60/673,599
; PRIOR FILING DATE: 2005-04-20
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 300
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Escherichia coli
PCT-US06-15100-300
```

## Alignment Scores:

Pred. No.:	2.04	Length:	2814
Score:	98.00	Matches:	79
Percent Similarity:	33.1%	Conservative:	53
Best Local Similarity:	19.8%	Mismatches:	123
Query Match:	6.3%	Indels:	144
DB:	1	Gaps:	20

US-10-734-719-14 (1-291) x PCT-US06-15100-300 (1-2814)

```
QY 5 IleIleAlaGlyAsnGlyProSerLeu----- 13
Db 418 ATTTTAACGGGGGAATGGACCATCTATATTTTAAACGGAAAAACAATAAACACATGAT 477
QY 14 -----LysGluIleAspTyrSerArgLeuProAsn-----Asp 24
Db 478 GGAGCTTTTGGAAAAATAATAGATTTTACCACAAATAAATAAATAAATAAAGTGCACAAAGAT 537
QY 25 PheAspValPheArgCysAsnGlnPheTyrPheGlu-----AspLys 38
Db 538 AAGAAATAATAAAA-----AGATATATTTTGTGAAGATCAATGTTTGTGATTAAGAG 591
QY 39 TyrTyrLeu-----GlyLysLysCysLysThrValPhe-----TyrThrProAsnPhe 54
Db 592 TATTTCACTAGAACTCGTAAAAAATTTCTCATACCCCATTTGAAAGTTTTTATAAAGAGCGA 651
QY 55 PhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnAsn----- 68
Db 652 TTATTTAATCAATATGATATATTTTAAACAACTTGCTAGAGATATATTGGGCACACCCATTA 711
QY 69 -----GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyr 81
Db 712 GTAATATATGATCGGGTTACAAAAATAATCATGAGTCATTAATGCTGCTATTGCAATA 771
QY 82 AsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro--- 100
Db 772 AACTCTAGACCAATTACAAAGAAATACATATATGGTGATGTTTATATCATTAATAAAAAA 831
QY 101 AspAlaHisLeuGlyTyrAsp-----PhePheLysGluLeuLys 113
Db 832 GACATCTCTTTGGAGTAGATACTCTTTACGGGAGGAAAGTTTGTATATTGTACTGGAT 891
QY 114 GluPheAsnAlaTyr----- 118
Db 892 GCAATGACGTATATAGAAAAAGCAAAAAAATGAGAGTTATTTTCCAATAATGAGATGAAA 951
QY 119 -----PheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMet 136
Db 952 AAAAGCATTAATAATATCTGAATAGTTATCCATAATATATAAAGAAAGGATGACTAAT 1011
QY 137 CysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr 156
Db 1012 TGTTCGCTAAA-----AAGGATCTTCTTAATGCCAGATATGATCTTATT 1056
QY 157 GlnAsnGlySerSerTyrAlaPheAspThrLysGln----- 168
Db 1057 AGAGATATTTCTCGATATTTCTTAAATATACGACAGGGAATTAACATGATGATGTTAAT 1116
QY 169 -----GluAsnLeuLeuLys-----LeuAlaProAspPhe 178
Db 1117 AGAATAGCGGAAATAATAAATAAAGATGATGGTATACTGAGGGGTATGAATCCTTAAACCT 1176
```

```
Qy 179 LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGlu 198
Db 1177 AGGAATGCCAGATA-----TCTAAAGAATTGCTTTATTAGCTGTTGAT 1221
Qy 199 -----PheLeuGluLysThrTyrLysIleLysLeuTyr 209
Db 1222 AGACAGATTGAGTGGCGGAAAAACATTTTATACAAAAGATGTATTGGAATGTTGTG 1281
Qy 210 CysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSer 229
Db 1282 TCAAAATGTGATTATCATCTATCTTTAATGTTAATAAAGTGCTTCAGAATACTATTCTT 1341
Qy 230 AsnPheIleIleGln-----GluLys 236
Db 1342 GAGTTTGTCATGAAATTAATAATATATCATCTCTCGCTGGATGTCAAAATCAGAAAAG 1401
Qy 237 AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLys 256
Db 1402 AATAATAACAAAA-----GAGGCAATAGAAAAAGTTCAAAAAA 1440
Qy 257 AsnIleAsn-----PheLysLysIleLysIle 265
Db 1441 GAAGTATCCCATATGAATGGCGGCAGAGTTTATTTGGGGGTTTGATAAGGTT---ATT 1497
Qy 266 LysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIle 284
Db 1498 CAAGAGGC-----TTAAGTGGATTGATTGAGTTAAGTATCGATATT 1539
```

Search completed: August 21, 2006, 17:16:39  
Job time : 1100 secs

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: August 21, 2006, 17:07:40 ; Search time 213 Seconds  
(without alignments)  
7695.264 Million cell updates/sec

Title: US-10-734-719-8  
Perfect score: 876  
Sequence: 1 atgaaaaagtattattgc.....attattcaaggaaataaa 876

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents NA:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:  
2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:  
3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq:  
4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq:  
5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:  
6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq:  
7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCUTUS COMB.seq:  
8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP COMB.seq:  
9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq:  
10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	876	100.0	876	3	US-09-495-406-8
2	876	100.0	876	3	US-09-816-028A-8
3	876	100.0	876	3	US-10-303-162-8
4	876	100.0	876	3	US-10-303-134-8
5	876	100.0	876	3	US-10-303-118-8
6	876	100.0	876	3	US-10-303-128-8
7	876	100.0	876	5	US-10-735-419-8
8	869.8	99.3	873	3	US-09-816-028A-13
9	869.8	99.3	873	3	US-10-303-162-13
10	869.8	99.3	873	3	US-10-303-134-13
11	869.8	99.3	873	3	US-10-303-118-13
12	869.8	99.3	873	5	US-10-303-128-13
13	869.8	99.3	873	5	US-10-735-419-13
14	866.6	98.9	873	3	US-09-816-028A-11
15	866.6	98.9	873	3	US-10-303-162-11
16	866.6	98.9	873	3	US-10-303-134-11
17	866.6	98.9	873	3	US-10-303-118-11
18	866.6	98.9	873	5	US-10-303-128-11
19	866.6	98.9	873	5	US-10-735-419-11
20	844	96.3	876	3	US-09-495-406-4
21	844	96.3	876	3	US-09-816-028A-4
22	844	96.3	876	3	US-10-303-162-4
23	844	96.3	876	3	US-10-303-134-4

96.3	876	3	US-10-303-118-4	Sequence 4, Appli
96.3	876	3	US-10-303-128-4	Sequence 4, Appli
96.3	876	5	US-10-735-419-4	Sequence 4, Appli
95.6	11474	3	US-09-495-406-1	Sequence 1, Appli
95.6	11474	3	US-09-816-028A-1	Sequence 1, Appli
95.6	11474	3	US-10-303-162-1	Sequence 1, Appli
95.6	11474	3	US-10-303-134-1	Sequence 1, Appli
95.6	11474	3	US-10-303-118-1	Sequence 1, Appli
95.6	11474	3	US-10-303-128-1	Sequence 1, Appli
95.6	11474	5	US-10-735-419-1	Sequence 1, Appli
95.4	876	3	US-09-495-406-2	Sequence 2, Appli
95.4	876	3	US-09-816-028A-2	Sequence 2, Appli
95.4	876	3	US-10-303-162-2	Sequence 2, Appli
95.4	876	3	US-10-303-134-2	Sequence 2, Appli
95.4	876	3	US-10-303-118-2	Sequence 2, Appli
95.4	876	3	US-10-303-128-2	Sequence 2, Appli
95.4	876	5	US-10-735-419-2	Sequence 2, Appli
93.6	876	3	US-09-495-406-6	Sequence 6, Appli
93.6	876	3	US-09-816-028A-6	Sequence 6, Appli
93.6	876	3	US-10-303-162-6	Sequence 6, Appli
93.6	876	3	US-10-303-134-6	Sequence 6, Appli
93.6	876	3	US-10-303-118-6	Sequence 6, Appli
93.6	876	3	US-10-303-128-6	Sequence 6, Appli
93.6	876	5	US-10-735-419-6	Sequence 6, Appli
35.6	1293	3	US-09-272-960-1	Sequence 1, Appli
35.6	1293	3	US-10-058-636-1	Sequence 1, Appli
32.7	1830121	3	US-09-557-884-1	Sequence 1, Appli
32.7	1830121	3	US-09-643-990A-1	Sequence 1, Appli
32.7	1830121	3	US-10-158-865-1	Sequence 1, Appli
9.0	205044	3	US-09-949-016-15851	Sequence 15851, A
9.0	205044	3	US-09-949-016-15852	Sequence 15852, A
9.0	205044	3	US-09-949-016-15853	Sequence 15853, A
9.0	223471	3	US-09-949-016-12387	Sequence 12387, A
9.0	223471	3	US-09-949-016-12387	Sequence 12387, A
9.0	223471	3	US-09-949-016-12724	Sequence 12724, A
9.0	223471	3	US-09-949-016-12725	Sequence 12725, A
7.9	187169	3	US-09-949-016-12776	Sequence 12776, A
7.9	191569	3	US-09-949-016-15940	Sequence 15940, A
7.9	25590	3	US-09-949-002-777	Sequence 777, App
7.6	601	3	US-09-949-016-30531	Sequence 30531, A
7.6	601	3	US-09-949-016-37150	Sequence 37150, A
7.6	601	3	US-09-949-016-37164	Sequence 37164, A
7.6	601	3	US-09-949-016-145868	Sequence 145868, A
7.6	601	3	US-09-949-016-146136	Sequence 146136, A
7.6	601	3	US-09-949-016-146404	Sequence 146404, A
7.5	1055	3	US-09-806-708B-23	Sequence 23, Appl
7.5	6243	2	US-09-056-075-1	Sequence 1, Appli
7.4	231129	3	US-09-949-016-16110	Sequence 16110, A
7.4	266293	3	US-09-949-016-11934	Sequence 11934, A
7.4	134987	3	US-09-949-016-15348	Sequence 15348, A
7.4	134987	3	US-09-949-016-15349	Sequence 15349, A
7.4	134987	3	US-09-949-016-15350	Sequence 15350, A
7.4	134987	3	US-09-949-016-15507	Sequence 15507, A
7.4	134987	3	US-09-949-016-15507	Sequence 15507, A
7.4	134987	3	US-09-949-016-15508	Sequence 15508, A
7.4	134987	3	US-09-949-016-15509	Sequence 15509, A
7.4	50000	3	US-09-662-254B-23	Sequence 23, Appl
7.3	612	3	US-09-902-540-1357	Sequence 1357, Ap
7.3	1141	3	US-09-806-708B-22	Sequence 22, Appl
7.3	32392	3	US-09-662-254B-27	Sequence 27, Appl
7.3	134987	3	US-09-949-016-15348	Sequence 15348, A
7.3	134987	3	US-09-949-016-15349	Sequence 15349, A
7.3	134987	3	US-09-949-016-15350	Sequence 15350, A
7.3	134987	3	US-09-949-016-15507	Sequence 15507, A
7.3	134987	3	US-09-949-016-15508	Sequence 15508, A
7.3	134987	3	US-09-949-016-15509	Sequence 15509, A
7.2	50000	3	US-09-662-254B-25	Sequence 25, Appl
7.2	50383	3	US-09-949-016-17600	Sequence 17600, A
7.2	50383	3	US-09-949-016-17600	Sequence 17600, A
7.1	1039	3	US-09-902-540-1280	Sequence 1280, Ap
7.1	18773	3	US-09-949-016-14164	Sequence 14164, A
7.0	19438	3	US-09-949-016-12699	Sequence 12699, A
7.0	1141	3	US-09-806-708B-22	Sequence 22, Appl
7.0	601	3	US-09-949-016-156535	Sequence 156535, A
7.0	50000	3	US-09-662-254B-24	Sequence 24, Appl

C 97	60.8	6.9	601	3	US-09-949-016-30532	Sequence 30532, A	C 170	54.4	6.2	640681	3	US-09-790-988-1	Sequence 1, Appli
C 98	60.8	6.9	601	3	US-09-949-016-37151	Sequence 37151, A	C 171	54.2	6.2	18651	3	US-09-949-002-592	Sequence 592, App
C 99	60.8	6.9	601	3	US-09-949-016-37165	Sequence 37165, A	C 172	54.2	6.2	18682	3	US-09-949-002-786	Sequence 786, App
C 100	60.8	6.9	601	3	US-09-949-016-145869	Sequence 145869, A	C 173	54.2	6.2	20674	3	US-09-641-638-651	Sequence 651, App
C 101	60.8	6.9	601	3	US-09-949-016-146137	Sequence 146137, A	C 174	54.2	6.2	20674	3	US-10-170-097-651	Sequence 251, App
C 102	60.8	6.9	601	3	US-09-949-016-146405	Sequence 146405, A	C 175	54.2	6.2	50000	3	US-09-662-254B-24	Sequence 24, Appl
C 103	60.8	6.9	32392	3	US-09-662-254B-27	Sequence 27, Appl	C 176	54	6.2	601	3	US-09-949-016-196366	Sequence 196366, A
C 104	60.4	6.9	601	3	US-09-949-016-186037	Sequence 186037, A	C 177	54	6.2	601	3	US-09-949-002-7549	Sequence 7549, Ap
C 105	60.4	6.9	45183	3	US-09-949-016-12798	Sequence 12798, A	C 178	54	6.2	601	3	US-09-710-279-997	Sequence 997, App
C 106	60.4	6.9	96922	3	US-09-949-016-17061	Sequence 17061, A	C 179	54	6.2	852	3	US-09-134-001C-786	Sequence 786, App
C 107	60.2	6.9	19124	2	US-08-487-826B-13	Sequence 13, Appl	C 180	54	6.2	3089	3	US-09-710-279-3965	Sequence 3965, Ap
C 108	59.8	6.8	6124	3	US-08-213-419B-3	Sequence 3, Appli	C 181	54	6.2	55886	3	US-09-949-016-15129	Sequence 15129, A
C 109	59.8	6.8	50000	3	US-09-662-254B-25	Sequence 25, Appl	C 182	54	6.2	94142	3	US-09-949-016-16553	Sequence 16553, A
C 110	59.4	6.8	50000	3	US-09-662-254B-23	Sequence 23, Appl	C 183	53.8	6.1	50000	3	US-09-662-254B-26	Sequence 26, Appl
C 111	59.4	6.8	194537	3	US-09-949-016-12928	Sequence 12928, A	C 184	53.6	6.1	601	3	US-09-949-002-7548	Sequence 7548, Ap
C 112	59.4	6.8	201529	3	US-09-949-016-12740	Sequence 12740, A	C 185	53.6	6.1	30820	3	US-09-949-016-17145	Sequence 17145, A
C 113	59	6.7	19124	2	US-08-487-826B-13	Sequence 13, Appl	C 186	53.2	6.1	601	3	US-09-949-016-196365	Sequence 196365, A
C 114	58.8	6.7	95255	3	US-09-949-016-17067	Sequence 17067, A	C 187	52.8	6.0	601	3	US-09-949-016-132806	Sequence 132806, A
C 115	58.2	6.6	50000	3	US-09-662-254B-26	Sequence 26, Appl	C 188	52.8	6.0	1431	3	US-09-316-083-2	Sequence 2, Appli
C 116	58.2	6.6	263893	3	US-09-949-016-12386	Sequence 12386, A	C 189	52.8	6.0	1431	3	US-09-933-700-2	Sequence 2, Appli
C 117	58.2	6.6	263894	3	US-09-949-016-16915	Sequence 16915, A	C 190	52.8	6.0	1830	3	US-09-662-254B-67	Sequence 67, Appl
C 118	58	6.6	55264	3	US-09-949-016-15014	Sequence 15014, A	C 191	52.8	6.0	194714	3	US-09-949-016-11869	Sequence 11869, A
C 119	58	6.6	94755	3	US-09-949-016-11839	Sequence 11839, A	C 192	52.8	6.0	196714	3	US-09-949-016-15474	Sequence 15474, A
C 120	57.8	6.6	601	3	US-09-949-016-179399	Sequence 179399, A	C 193	52.6	6.0	615	3	US-08-998-416-186	Sequence 186, App
C 121	57.6	6.6	105001	3	US-09-949-002-627	Sequence 627, App	C 194	52.6	6.0	114139	3	US-09-949-016-16536	Sequence 16536, A
C 122	57.6	6.6	105002	3	US-09-949-002-821	Sequence 821, App	C 195	52.6	6.0	205044	3	US-09-949-016-15851	Sequence 15851, A
C 123	57.4	6.6	1132	3	US-09-533-559-5452	Sequence 5452, Ap	C 196	52.6	6.0	205044	3	US-09-949-016-15852	Sequence 15852, A
C 124	57.4	6.6	67755	3	US-09-949-016-13703	Sequence 13703, A	C 197	52.6	6.0	205044	3	US-09-949-016-15853	Sequence 15853, A
C 125	56.8	6.5	6124	3	US-08-213-419B-3	Sequence 3, Appli	C 198	52.6	6.0	223471	3	US-09-949-016-12387	Sequence 12387, A
C 126	56.6	6.5	861	3	US-09-662-254B-1	Sequence 1, Appl	C 199	52.6	6.0	223471	3	US-09-949-016-12724	Sequence 12724, A
C 127	56.4	6.4	45587	3	US-09-949-016-15836	Sequence 15836, A	C 200	52.6	6.0	223471	3	US-09-949-016-12725	Sequence 12725, A
C 128	56.2	6.4	601	3	US-09-949-016-15836	Sequence 30533, A	C 201	52.4	6.0	601	3	US-09-949-016-132805	Sequence 132805, A
C 129	56.2	6.4	601	3	US-09-949-016-30533	Sequence 30534, A	C 202	52.4	6.0	3701	3	US-08-845-258-10	Sequence 10, Appl
C 130	56.2	6.4	601	3	US-09-949-016-37152	Sequence 37152, A	C 203	52.4	6.0	3701	3	US-08-990-571-10	Sequence 10, Appl
C 131	56.2	6.4	601	3	US-09-949-016-37153	Sequence 37153, A	C 204	52.4	6.0	3701	3	US-08-723-142A-10	Sequence 10, Appl
C 132	56.2	6.4	601	3	US-09-949-016-37166	Sequence 37166, A	C 205	52.4	6.0	3701	3	US-09-528-784A-10	Sequence 10, Appl
C 133	56.2	6.4	601	3	US-09-949-016-37167	Sequence 37167, A	C 206	52.4	6.0	3701	3	US-09-569-098A-10	Sequence 10, Appl
C 134	56.2	6.4	601	3	US-09-949-016-145870	Sequence 145870, A	C 207	52.4	6.0	55886	3	US-09-949-016-15129	Sequence 15129, A
C 135	56.2	6.4	601	3	US-09-949-016-145871	Sequence 145871, A	C 208	52.4	6.0	119153	3	US-09-949-016-12378	Sequence 12378, A
C 136	56.2	6.4	601	3	US-09-949-016-146138	Sequence 146138, A	C 209	52.4	6.0	1664976	3	US-08-916-421B-1	Sequence 1, Appli
C 137	56.2	6.4	601	3	US-09-949-016-146139	Sequence 146139, A	C 210	52.4	6.0	1664976	3	US-09-692-570-1	Sequence 1, Appli
C 138	56.2	6.4	601	3	US-09-949-016-146406	Sequence 146406, A	C 211	52.2	6.0	263693	3	US-09-949-016-12386	Sequence 12386, A
C 139	56.2	6.4	601	3	US-09-949-016-146407	Sequence 146407, A	C 212	52.2	6.0	263694	3	US-09-949-016-16915	Sequence 16915, A
C 140	56	6.4	837	3	US-08-998-416-288	Sequence 288, App	C 213	52	5.9	601	3	US-09-949-016-132807	Sequence 132807, A
C 141	55.8	6.4	53332	3	US-09-801-861-3	Sequence 3, Appli	C 214	52	5.9	636	3	US-08-998-416-1137	Sequence 1137, Ap
C 142	55.8	6.4	53332	3	US-10-224-562-3	Sequence 3, Appli	C 215	52	5.9	3312	3	US-09-601-198-58	Sequence 58, Appl
C 143	55.8	6.4	53332	3	US-10-786-065-3	Sequence 3, Appli	C 216	52	5.9	108310	3	US-09-949-016-16366	Sequence 16366, A
C 144	55.6	6.3	96922	3	US-09-949-016-61868	Sequence 17061, A	C 217	51.8	5.9	12313	3	US-09-949-016-13248	Sequence 13248, A
C 145	55.4	6.3	582	3	US-09-949-016-61868	Sequence 61868, A	C 218	51.8	5.9	25067	3	US-09-949-016-11794	Sequence 11794, A
C 146	55.4	6.3	314798	3	US-09-949-016-13539	Sequence 13539, A	C 219	51.8	5.9	25441	3	US-09-949-016-14232	Sequence 14232, A
C 147	55.2	6.3	601	3	US-09-949-002-7550	Sequence 7550, Ap	C 220	51.8	5.9	66933	3	US-09-544-398B-11	Sequence 11, Appl
C 148	55.2	6.3	700	3	US-09-735-271-1038	Sequence 1038, Ap	C 221	51.8	5.9	66933	3	US-09-543-771B-11	Sequence 11, Appl
C 149	55.2	6.3	137226	3	US-09-949-016-13763	Sequence 13763, A	C 222	51.8	5.9	72049	3	US-09-544-398B-9	Sequence 9, Appli
C 150	55.2	6.3	298336	3	US-09-949-016-16600	Sequence 16600, A	C 223	51.8	5.9	72049	3	US-09-543-771B-9	Sequence 9, Appli
C 151	55.2	6.3	451924	3	US-09-949-016-12896	Sequence 12896, A	C 224	51.8	5.9	140844	3	US-09-949-016-14199	Sequence 14199, A
C 152	55.2	6.3	451925	3	US-09-949-016-17305	Sequence 17305, A	C 225	51.6	5.9	266223	3	US-09-949-016-16110	Sequence 16110, A
C 153	55.2	6.3	640681	3	US-09-790-988-1	Sequence 1, Appli	C 226	51.6	5.9	266223	3	US-09-949-016-11934	Sequence 11934, A
C 154	55	6.3	601	3	US-09-949-016-30530	Sequence 30530, A	C 227	51.4	5.9	396	3	US-09-640-173-53	Sequence 53, Appl
C 155	55	6.3	601	3	US-09-949-016-37149	Sequence 37149, A	C 228	51.4	5.9	396	3	US-09-713-550-53	Sequence 53, Appl
C 156	55	6.3	601	3	US-09-949-016-37163	Sequence 37163, A	C 229	51.4	5.9	396	3	US-09-825-294-53	Sequence 53, Appl
C 157	55	6.3	601	3	US-09-949-016-145867	Sequence 145867, A	C 230	51.4	5.9	396	3	US-09-970-966-53	Sequence 53, Appl
C 158	55	6.3	601	3	US-09-949-016-146135	Sequence 146135, A	C 231	51.4	5.9	601	3	US-09-949-016-87162	Sequence 87162, A
C 159	55	6.3	601	3	US-09-949-016-146403	Sequence 146403, A	C 232	51.4	5.9	5340	3	US-09-627-122-21	Sequence 21, Appl
C 160	55	6.3	19438	3	US-09-949-016-12699	Sequence 12699, A	C 233	51.4	5.9	10764	3	US-09-949-016-11872	Sequence 11872, A
C 161	55	6.3	119153	3	US-09-949-016-12378	Sequence 12378, A	C 234	51.4	5.9	10765	3	US-09-949-016-14204	Sequence 14204, A
C 162	55	6.3	237241	3	US-09-949-016-16101	Sequence 16101, A	C 235	51.4	5.9	42252	3	US-09-949-016-14182	Sequence 14182, A
C 163	54.8	6.3	18773	3	US-09-949-016-14164	Sequence 14164, A	C 236	51.2	5.8	658	3	US-08-998-416-595	Sequence 595, App
C 164	54.6	6.2	1500	3	US-09-601-198-36	Sequence 36, Appl	C 237	51.2	5.8	2230	3	US-08-844-188-34	Sequence 34, Appl
C 165	54.4	6.2	1316	3	US-09-270-767-2481	Sequence 2481, Ap	C 238	51.2	5.8	2230	3	US-09-378-088A-34	Sequence 34, Appl
C 166	54.4	6.2	1316	3	US-09-270-767-17763	Sequence 17763, A	C 239	51.2	5.8	2230	3	US-09-548-334A-34	Sequence 34, Appl
C 167	54.4	6.2	25590	3	US-09-949-002-777	Sequence 777, App	C 240	51.2	5.8	2230	3	US-09-547-621-34	Sequence 34, Appl
C 168	54.4	6.2	55264	3	US-09-949-016-15014	Sequence 15014, A	C 241	51.2	5.8	2230	3	US-09-643-596B-34	Sequence 34, Appl
C 169	54.4	6.2	94755	3	US-09-949-016-11839	Sequence 11839, A	C 242	51.2	5.8	2230	3	US-10-412-203-34	Sequence 34, Appl

c 243	51.2	5.8	2230	3	US-10-099-278-34	Sequence 34, Appl	316	48.2	5.5	601	3	US-09-949-016-146135	Sequence 146135,
c 244	51.2	5.8	285986	3	US-09-949-016-12287	Sequence 12287, A	317	48.2	5.5	601	3	US-09-949-016-146403	Sequence 146403,
c 245	51.2	5.8	288031	3	US-09-949-016-14864	Sequence 14864, A	318	48.2	5.5	605	3	US-08-358-918-36	Sequence 36, Appl
c 246	51	5.8	1885	3	US-09-662-254B-41	Sequence 41, Appl	c 319	48.2	5.5	665	2	US-08-883-795A-36	Sequence 36, Appl
c 247	51	5.8	6508	3	US-09-995-917A-2	Sequence 2, Appl1	320	48.2	5.5	202001	3	US-09-734-674-3	Sequence 3, Appl1
c 248	51	5.8	95255	3	US-09-949-016-17067	Sequence 17067, A	321	48.2	5.5	202001	3	US-10-274-990-3	Sequence 211, App
c 249	50.8	5.8	4529	3	US-09-949-016-14004	Sequence 14004, A	c 322	48.2	5.5	1082144	4	US-09-531-120-211	Sequence 237, App
c 250	50.8	5.8	29604	3	US-08-781-891-207	Sequence 207, App	c 323	48	5.5	591	3	US-09-702-705-237	Sequence 237, App
c 251	50.8	5.8	29604	3	US-09-618-166-207	Sequence 207, App	c 324	48	5.5	591	3	US-09-736-457-237	Sequence 237, App
c 252	50.8	5.8	134537	3	US-09-949-016-12928	Sequence 12928, A	c 325	48	5.5	591	3	US-09-614-124B-237	Sequence 237, App
c 253	50.8	5.8	201529	3	US-09-949-016-12740	Sequence 12740, A	c 326	48	5.5	591	3	US-09-671-325-237	Sequence 237, App
c 254	50.6	5.8	601	3	US-09-949-016-30531	Sequence 30531, A	c 327	48	5.5	591	3	US-09-589-184-237	Sequence 237, App
c 255	50.6	5.8	601	3	US-09-949-016-37150	Sequence 37150, A	c 328	48	5.5	591	3	US-09-658-824-237	Sequence 237, App
c 256	50.6	5.8	601	3	US-09-949-016-37164	Sequence 37164, A	c 329	48	5.5	591	3	US-10-017-754-237	Sequence 237, App
c 257	50.6	5.8	601	3	US-09-949-016-145868	Sequence 145868, A	c 330	48	5.5	591	3	US-09-651-563-237	Sequence 237, App
c 258	50.6	5.8	601	3	US-09-949-016-146136	Sequence 146136, A	c 331	48	5.5	591	3	US-09-519-642-237	Sequence 237, App
c 259	50.6	5.8	601	3	US-09-949-016-146404	Sequence 146404, A	c 332	48	5.5	601	3	US-09-949-016-86305	Sequence 86305, A
c 260	50.6	5.8	251672	3	US-09-949-016-17296	Sequence 17296, A	c 333	48	5.5	2132	3	US-08-844-188-39	Sequence 39, Appl
c 261	50.6	5.8	251682	3	US-09-949-016-17296	Sequence 17296, A	c 334	48	5.5	2132	3	US-09-378-088A-39	Sequence 39, Appl
c 262	50.4	5.8	3591	3	US-09-662-254B-35	Sequence 35, Appl	c 335	48	5.5	2132	3	US-09-548-334A-39	Sequence 39, Appl
c 263	50.2	5.7	50368	3	US-09-949-016-13256	Sequence 13256, A	c 336	48	5.5	2132	3	US-09-547-621-39	Sequence 39, Appl
c 264	50.2	5.7	61178	3	US-09-949-016-17369	Sequence 17369, A	c 337	48	5.5	2132	3	US-09-643-566B-39	Sequence 39, Appl
c 265	50	5.7	1458	3	US-09-830-230A-100	Sequence 100, App	c 338	48	5.5	2132	3	US-10-412-203-39	Sequence 39, Appl
c 266	50	5.7	1521	3	US-09-830-230A-99	Sequence 99, Appl	c 339	48	5.5	2132	3	US-10-099-278-39	Sequence 39, Appl
c 267	50	5.7	1866	3	US-09-601-198-153	Sequence 153, App	c 340	48	5.5	5613	2	US-08-463-418-1	Sequence 1, Appl1
c 268	50	5.7	62957	3	US-09-949-016-12676	Sequence 12676, A	c 341	48	5.5	5852	2	US-07-867-106-2	Sequence 2, Appl1
c 269	50	5.7	63714	3	US-09-949-016-16289	Sequence 16289, A	c 342	48	5.5	9955	4	US-09-747-385-15	Sequence 15, Appl
c 270	50	5.7	147382	3	US-09-949-016-14624	Sequence 14624, A	c 343	48	5.5	9955	4	US-09-949-016-13504	Sequence 13504, A
c 271	49.8	5.7	3095	10	5231168-1	Patent No. 5231168	c 344	48	5.5	213117	4	US-09-531-120-212	Sequence 212, App
c 272	49.8	5.7	114139	3	US-09-949-016-16536	Sequence 16536, A	c 345	48	5.5	111049	3	US-09-949-016-15770	Sequence 15770, A
c 273	49.8	5.7	253375	3	US-09-949-016-12849	Sequence 12849, A	c 346	47.8	5.5	601	3	US-09-949-016-28525	Sequence 28525, A
c 274	49.6	5.7	1407	3	US-09-662-254B-50	Sequence 50, Appl	c 347	47.8	5.5	601	3	US-09-949-016-28526	Sequence 28526, A
c 275	49.6	5.7	2427	3	US-09-601-198-70	Sequence 70, Appl	c 348	47.8	5.5	601	3	US-09-949-016-60958	Sequence 60958, A
c 276	49.6	5.7	30820	3	US-09-949-016-17145	Sequence 17145, A	c 349	47.8	5.5	601	3	US-09-949-016-60959	Sequence 60959, A
c 277	49.4	5.6	601	3	US-09-949-016-28524	Sequence 28524, A	c 350	47.8	5.5	3255	3	US-09-601-198-108	Sequence 108, App
c 278	49.4	5.6	2181	3	US-09-949-016-60957	Sequence 60957, A	c 351	47.8	5.5	29171	3	US-09-949-016-12283	Sequence 12283, A
c 279	49.4	5.6	6881	3	US-09-662-254B-40	Sequence 40, Appl	c 352	47.8	5.5	29171	3	US-09-949-016-13509	Sequence 13509, A
c 280	49.4	5.6	6881	3	US-09-949-016-17181	Sequence 17181, A	c 353	47.8	5.5	57507	3	US-09-949-016-15019	Sequence 15019, A
c 281	49.4	5.6	84495	3	US-09-797-906-3	Sequence 3, Appl1	c 354	47.8	5.5	84252	3	US-09-949-016-17315	Sequence 17315, A
c 282	49.2	5.6	614	3	US-09-902-540-1318	Sequence 1318, Ap	c 355	47.8	5.5	251672	3	US-09-949-016-17296	Sequence 17296, A
c 283	49.2	5.6	702	3	US-09-543-681A-4167	Sequence 4167, Ap	c 356	47.8	5.5	251682	3	US-09-949-016-11973	Sequence 11973, A
c 284	49.2	5.6	15016	3	US-09-601-198-60	Sequence 60, Appl	c 357	47.6	5.4	688	3	US-08-998-416-972	Sequence 972, App
c 285	49.2	5.6	133719	3	US-09-949-016-15092	Sequence 15092, A	c 358	47.6	5.4	831	3	US-09-662-254B-69	Sequence 69, Appl
c 286	49.2	5.6	163022	3	US-09-949-016-16515	Sequence 16515, A	c 359	47.6	5.4	861	3	US-09-601-198-66	Sequence 66, Appl
c 287	49.2	5.6	422592	3	US-09-949-016-14182	Sequence 14182, A	c 360	47.6	5.4	7358	3	US-09-058-483-8	Sequence 8, Appl1
c 288	49	5.6	700	3	US-09-735-271-1038	Sequence 70, Appl	c 361	47.6	5.4	7633	3	US-09-028-851-1	Sequence 1, Appl1
c 289	49	5.6	1002	3	US-09-662-254B-70	Sequence 70, Appl	c 362	47.6	5.4	7633	3	US-08-815-520-1	Sequence 1, Appl1
c 290	48.8	5.6	601	3	US-09-949-016-68706	Sequence 68706, A	c 363	47.6	5.4	7633	3	US-09-273-163-1	Sequence 1, Appl1
c 291	48.8	5.6	601	3	US-09-949-016-186365	Sequence 186365, A	c 364	47.6	5.4	67620	3	US-09-949-016-16939	Sequence 16939, A
c 292	48.8	5.6	701	3	US-08-998-416-701	Sequence 701, App	c 365	47.6	5.4	15841	3	US-09-949-016-16192	Sequence 16192, A
c 293	48.8	5.6	25199	3	US-09-949-016-13361	Sequence 13361, A	c 366	47.6	5.4	298336	3	US-09-949-016-16600	Sequence 16600, A
c 294	48.8	5.6	32039	3	US-09-949-016-16647	Sequence 16647, A	c 367	47.4	5.4	601	3	US-09-949-016-51985	Sequence 51985, A
c 295	48.8	5.6	32043	3	US-09-949-016-16647	Sequence 16647, A	c 368	47.4	5.4	601	3	US-09-949-016-51986	Sequence 51986, A
c 296	48.6	5.5	665	2	US-08-883-795A-36	Sequence 36, Appl	c 369	47.4	5.4	605	2	US-08-358-918-36	Sequence 36, Appl
c 297	48.6	5.5	1830	3	US-09-662-254B-67	Sequence 67, Appl	c 370	47.4	5.4	18651	3	US-09-949-002-592	Sequence 592, App
c 298	48.6	5.5	2486	5	US-09-497-967-2	Sequence 2, Appl1	c 371	47.4	5.4	18682	3	US-09-949-002-786	Sequence 786, App
c 299	48.6	5.5	10640	5	US-09-417-485D-5	Sequence 5, Appl1	c 372	47.4	5.4	27916	3	US-09-949-016-15202	Sequence 15202, A
c 300	48.6	5.5	17082	3	US-09-949-016-14893	Sequence 14893, A	c 373	47.4	5.4	105055	3	US-09-949-016-14001	Sequence 14001, A
c 301	48.6	5.5	29717	3	US-09-949-016-16284	Sequence 16284, A	c 374	47.4	5.4	387902	3	US-09-949-016-14543	Sequence 14543, A
c 302	48.6	5.5	60376	3	US-09-949-016-12423	Sequence 12423, A	c 375	47.4	5.4	421883	3	US-09-949-016-12557	Sequence 12557, A
c 303	48.4	5.5	601	3	US-09-949-016-156535	Sequence 156535, A	c 376	47.2	5.4	700	3	US-09-735-271-1039	Sequence 1039, Ap
c 304	48.4	5.5	601	3	US-09-949-016-170363	Sequence 170363, A	c 377	47.2	5.4	729	3	US-09-601-198-16	Sequence 16, Appl
c 305	48.4	5.5	1308	3	US-09-601-198-3	Sequence 3, Appl1	c 378	47.2	5.4	1170	3	US-09-495-406-11	Sequence 11, Appl
c 306	48.4	5.5	9447	3	US-09-949-016-14649	Sequence 14649, A	c 379	47.2	5.4	1170	3	US-09-816-028A-15	Sequence 15, Appl
c 307	48.4	5.5	10640	3	US-09-417-485D-5	Sequence 5, Appl1	c 380	47.2	5.4	1170	3	US-10-303-162-15	Sequence 15, Appl
c 308	48.4	5.5	168575	3	US-09-426-290-1	Sequence 1, Appl1	c 381	47.2	5.4	1170	3	US-10-303-134-15	Sequence 15, Appl
c 309	48.2	5.5	601	3	US-09-949-016-30530	Sequence 30530, A	c 382	47.2	5.4	1170	3	US-10-303-118-15	Sequence 15, Appl
c 310	48.2	5.5	601	3	US-09-949-016-37149	Sequence 37149, A	c 383	47.2	5.4	1170	5	US-10-303-128-15	Sequence 15, Appl
c 311	48.2	5.5	601	3	US-09-949-016-37163	Sequence 37163, A	c 384	47.2	5.4	6243	2	US-09-056-075-1	Sequence 1, Appl1
c 312	48.2	5.5	601	3	US-09-949-016-54020	Sequence 54020, A	c 385	47.2	5.4	6881	3	US-09-949-016-17181	Sequence 17181, A
c 313	48.2	5.5	601	3	US-09-949-016-108654	Sequence 108654, A	c 386	47.2	5.4	23652	3	US-09-949-016-16368	Sequence 16368, A
c 314	48.2	5.5	601	3	US-09-949-016-145867	Sequence 145867, A	c 387	47.2	5.4	67620	3	US-09-949-016-16939	Sequence 16939, A
c 315	48.2	5.5	601	3			c 388	47.2	5.4				



C 389	47.2	5.4	119930	3	US-09-949-016-12677	Sequence 12677, A	C 462	46.2	5.3	421118	3	US-09-949-016-16297	Sequence 16297, A
C 390	47.2	5.4	119931	3	US-09-949-016-16319	Sequence 16319, A	C 463	46	5.3	615	3	US-08-998-416-186	Sequence 186, App
C 391	47.2	5.4	125536	3	US-09-949-016-14186	Sequence 14186, A	C 464	46	5.3	731	2	US-08-451-405A-2	Sequence 2, Appli
C 392	47.2	5.4	129658	3	US-09-949-016-17195	Sequence 17195, A	C 465	46	5.3	2110	3	US-09-419-459-1	Sequence 1, Appli
C 393	47.2	5.4	139049	3	US-09-949-016-17030	Sequence 17030, A	C 466	46	5.3	6152	3	US-08-973-462-1	Sequence 1, Appli
C 394	47	5.4	601	3	US-09-949-016-15987	Sequence 15987, A	C 467	46	5.3	20674	3	US-09-641-638-651	Sequence 651, App
C 395	47	5.4	1055	3	US-09-806-708B-23	Sequence 23, Appl	C 468	46	5.3	20674	3	US-10-170-097-651	Sequence 937, App
C 396	47	5.4	47781	3	US-09-949-016-16492	Sequence 16492, A	C 469	45.8	5.2	665	3	US-08-998-416-937	Sequence 937, App
C 397	47	5.4	47781	3	US-09-949-016-16493	Sequence 16493, A	C 470	45.8	5.2	700	3	US-09-735-271-1039	Sequence 1039, Ap
C 398	47	5.4	47781	3	US-09-949-016-16494	Sequence 16494, A	C 471	45.8	5.2	1251	3	US-09-601-198-38	Sequence 38, Appl
C 399	46.8	5.3	453	3	US-09-662-254B-56	Sequence 56, Appl	C 472	45.8	5.2	1599	3	US-09-662-254B-58	Sequence 58, Appl
C 400	46.8	5.3	699	3	US-09-601-198-135	Sequence 135, App	C 473	45.8	5.2	2110	3	US-09-419-459-1	Sequence 1, Appli
C 401	46.8	5.3	2908	3	US-09-710-279-4359	Sequence 4359, A	C 474	45.8	5.2	2885	3	US-09-620-312D-900	Sequence 900, App
C 402	46.8	5.3	3160	3	US-09-710-279-4048	Sequence 4048, Ap	C 475	45.8	5.2	10478	3	US-09-445-774-16	Sequence 16, Appl
C 403	46.8	5.3	3706	3	US-08-913-159-9	Sequence 9, Appli	C 476	45.8	5.2	23501	3	US-09-949-016-17517	Sequence 17517, A
C 404	46.8	5.3	4344	3	US-09-601-198-165	Sequence 165, App	C 477	45.8	5.2	34011	3	US-09-949-016-12485	Sequence 12485, A
C 405	46.8	5.3	11122	5	US-09-984-429-394	Sequence 394, App	C 478	45.8	5.2	1664976	3	US-08-916-421B-1	Sequence 1, Appli
C 406	46.8	5.3	11122	5	US-09-984-429-444	Sequence 444, App	C 479	45.8	5.2	1664976	3	US-09-692-570-1	Sequence 5, Appli
C 407	46.8	5.3	46885	3	US-09-949-016-13848	Sequence 13848, A	C 480	45.6	5.2	1422	2	US-08-319-704-5	Sequence 32, Appl
C 408	46.8	5.3	137226	3	US-09-949-016-13763	Sequence 13763, A	C 481	45.6	5.2	1850	3	US-08-617-860B-32	Sequence 4780, Ap
C 409	46.8	5.3	317366	3	US-09-949-016-16001	Sequence 16001, A	C 482	45.6	5.2	1923	3	US-09-248-796A-4780	Sequence 2, Appli
C 410	46.8	5.3	387902	3	US-09-949-016-15453	Sequence 14543, A	C 483	45.6	5.2	2486	5	US-09-497-967-2	Sequence 2, Appli
C 411	46.8	5.3	421883	3	US-09-949-016-12557	Sequence 12557, A	C 484	45.6	5.2	2570	2	US-09-056-075-2	Sequence 2, Appli
C 412	46.8	5.3	1082144	4	US-09-531-120-211	Sequence 211, App	C 485	45.6	5.2	4098	2	US-08-605-106-4	Sequence 4, Appli
C 413	46.6	5.3	240	2	US-08-628-417-6	Sequence 6, Appli	C 486	45.6	5.2	5852	2	US-07-867-106-2	Sequence 2, Appli
C 414	46.6	5.3	601	3	US-09-949-016-54021	Sequence 54021, A	C 487	45.6	5.2	12313	3	US-09-949-016-13248	Sequence 13248, A
C 415	46.6	5.3	601	3	US-09-949-016-54024	Sequence 54024, A	C 488	45.6	5.2	29927	3	US-09-949-016-11814	Sequence 11814, A
C 416	46.6	5.3	601	3	US-09-949-016-108655	Sequence 108655, A	C 489	45.6	5.2	29927	3	US-09-949-016-17474	Sequence 17474, A
C 417	46.6	5.3	601	3	US-09-949-016-108658	Sequence 108658, A	C 490	45.6	5.2	29927	3	US-09-949-016-17475	Sequence 17475, A
C 418	46.6	5.3	1188	3	US-09-248-796A-11600	Sequence 11600, A	C 491	45.6	5.2	37802	3	US-09-949-016-12639	Sequence 12639, A
C 419	46.6	5.3	1431	3	US-09-316-083-2	Sequence 2, Appli	C 492	45.6	5.2	50186	3	US-09-949-016-14066	Sequence 14066, A
C 420	46.6	5.3	1431	3	US-09-933-700-2	Sequence 2, Appli	C 493	45.6	5.2	51967	3	US-09-949-016-16982	Sequence 16982, A
C 421	46.6	5.3	5046	3	US-09-725-735A-13	Sequence 13, Appl	C 494	45.6	5.2	67755	3	US-09-949-016-13703	Sequence 13703, A
C 422	46.6	5.3	42250	3	US-09-949-016-17171	Sequence 17171, A	C 495	45.6	5.2	8212	3	US-09-949-016-16823	Sequence 16823, A
C 423	46.6	5.3	101128	3	US-09-949-016-14293	Sequence 14293, A	C 496	45.4	5.2	601	3	US-09-949-016-196367	Sequence 196367, A
C 424	46.6	5.3	104520	3	US-09-949-016-13303	Sequence 13303, A	C 497	45.4	5.2	1299	3	US-09-710-279-779	Sequence 779, App
C 425	46.6	5.3	126029	3	US-09-949-016-14731	Sequence 14731, A	C 498	45.4	5.2	1719	3	US-09-662-254B-11	Sequence 11, Appl
C 426	46.6	5.3	264358	3	US-09-949-016-15725	Sequence 15725, A	C 499	45.4	5.2	3119	3	US-09-710-279-4270	Sequence 4270, Ap
C 427	46.4	5.3	601	3	US-09-949-002-3211	Sequence 3211, Ap	C 500	45.4	5.2	3399	3	US-09-601-198-43	Sequence 43, Appl
C 428	46.4	5.3	601	3	US-09-949-002-3212	Sequence 3212, Ap	C 501	45.4	5.2	3645	3	US-09-710-279-4110	Sequence 4110, Ap
C 429	46.4	5.3	601	3	US-09-949-002-8357	Sequence 8357, Ap	C 502	45.4	5.2	4152	3	US-09-662-254B-9	Sequence 9, Appli
C 430	46.4	5.3	601	3	US-09-949-002-8358	Sequence 8358, Ap	C 503	45.4	5.2	65415	3	US-09-949-016-16699	Sequence 16699, A
C 431	46.4	5.3	990	3	US-09-543-681A-1358	Sequence 1358, Ap	C 504	45.4	5.2	105413	3	US-10-427-923-3	Sequence 3, Appli
C 432	46.4	5.3	1407	3	US-09-662-254B-50	Sequence 50, Appl	C 505	45.4	5.2	112219	3	US-09-949-016-12453	Sequence 12453, A
C 433	46.4	5.3	3680	3	US-09-647-390-15	Sequence 15, Appl	C 506	45.4	5.2	112222	3	US-09-949-016-14324	Sequence 14324, A
C 434	46.4	5.3	123463	3	US-09-949-016-17078	Sequence 17078, A	C 507	45.4	5.2	113186	3	US-09-949-016-16192	Sequence 16192, A
C 435	46.4	5.3	193169	3	US-09-949-016-15091	Sequence 15091, A	C 508	45.4	5.2	165841	3	US-09-949-016-17572	Sequence 17572, A
C 436	46.2	5.3	518	3	US-09-123-913-87	Sequence 87, Appl	C 509	45.4	5.2	258775	3	US-09-949-016-13922	Sequence 16435, A
C 437	46.2	5.3	518	3	US-09-643-597-87	Sequence 87, Appl	C 510	45.4	5.2	278866	3	US-09-949-016-13922	Sequence 13922, A
C 438	46.2	5.3	518	3	US-09-480-884A-87	Sequence 87, Appl	C 511	45.4	5.2	278866	3	US-09-949-016-13923	Sequence 13923, A
C 439	46.2	5.3	518	3	US-09-542-615A-87	Sequence 87, Appl	C 512	45.4	5.2	278866	3	US-09-949-016-13924	Sequence 13924, A
C 440	46.2	5.3	518	3	US-09-606-421B-87	Sequence 87, Appl	C 513	45.4	5.2	278866	3	US-09-949-016-13925	Sequence 13925, A
C 441	46.2	5.3	518	3	US-09-221-107-87	Sequence 87, Appl	C 514	45.4	5.2	278866	3	US-09-949-016-13926	Sequence 13926, A
C 442	46.2	5.3	518	3	US-09-466-396A-87	Sequence 87, Appl	C 515	45.4	5.2	278866	3	US-09-949-016-14699	Sequence 14699, A
C 443	46.2	5.3	518	3	US-09-476-966A-87	Sequence 87, Appl	C 516	45.4	5.2	278866	3	US-09-949-016-14700	Sequence 14700, A
C 444	46.2	5.3	518	3	US-09-630-940B-87	Sequence 87, Appl	C 517	45.4	5.2	278866	3	US-09-949-016-14701	Sequence 14701, A
C 445	46.2	5.3	518	3	US-09-285-479-87	Sequence 87, Appl	C 518	45.4	5.2	278866	3	US-09-949-016-14702	Sequence 14702, A
C 446	46.2	5.3	518	3	US-10-007-700-87	Sequence 87, Appl	C 519	45.4	5.2	278866	3	US-09-949-016-14703	Sequence 14703, A
C 447	46.2	5.3	601	3	US-09-949-016-54022	Sequence 54022, A	C 520	45.4	5.2	636591	3	US-09-949-016-11808	Sequence 11808, A
C 448	46.2	5.3	601	3	US-09-949-016-54023	Sequence 54023, A	C 521	45.4	5.2	636591	3	US-09-949-016-13388	Sequence 13388, A
C 449	46.2	5.3	601	3	US-09-949-016-88182	Sequence 88182, A	C 522	45.2	5.2	601	3	US-09-949-016-86304	Sequence 86304, A
C 450	46.2	5.3	601	3	US-09-949-016-108656	Sequence 108656, A	C 523	45.2	5.2	601	3	US-09-949-002-3210	Sequence 3210, Ap
C 451	46.2	5.3	601	3	US-09-949-016-108657	Sequence 108657, A	C 524	45.2	5.2	601	3	US-09-949-002-8356	Sequence 8356, Ap
C 452	46.2	5.3	624	3	US-09-543-681A-985	Sequence 985, App	C 525	45.2	5.2	1267	4	US-09-297-648-4659	Sequence 4659, Ap
C 453	46.2	5.3	825	3	US-09-248-796A-9336	Sequence 9336, Ap	C 526	45.2	5.2	1511	2	US-07-991-867B-8	Sequence 8, Appli
C 454	46.2	5.3	832	3	US-09-621-975-2813	Sequence 2813, Ap	C 527	45.2	5.2	1511	2	US-08-107-755A-8	Sequence 8, Appli
C 455	46.2	5.3	1511	2	US-07-991-867B-8	Sequence 8, Appli	C 528	45.2	5.2	1511	2	US-08-544-332-8	Sequence 8, Appli
C 456	46.2	5.3	1511	2	US-08-107-755A-8	Sequence 8, Appli	C 529	45.2	5.2	2304	3	US-09-370-861A-8	Sequence 43, Appl
C 457	46.2	5.3	1511	2	US-08-544-332-8	Sequence 8, Appli	C 530	45.2	5.2	2304	3	US-09-662-254B-43	Sequence 13, Appl
C 458	46.2	5.3	1511	3	US-09-370-861A-8	Sequence 8, Appli	C 531	45.2	5.2	4148	3	US-09-435-019-13	Sequence 15, Appl
C 459	46.2	5.3	16550	3	US-08-916-421B-3	Sequence 3, Appli	C 532	45.2	5.2	4148	3	US-09-435-019-15	Sequence 15, Appl
C 460	46.2	5.3	16550	3	US-09-692-570-3	Sequence 3, Appli	C 533	45.2	5.2	4148	3	US-10-065-200A-13	Sequence 13, Appl
C 461	46.2	5.3	23790	3	US-09-949-016-14212	Sequence 14212, A	C 534	45.2	5.2	4148	3	US-10-065-200A-15	Sequence 15, Appl

C 535	45.2	5.2	14066	3	US-09-601-198-56	Sequence 56, Appl1	608	44.6	5.1	2476	3	US-10-012-064A-131	Sequence 131, App
C 536	45.2	5.2	18989	3	US-09-949-016-13435	Sequence 13435, A	609	44.6	5.1	2476	4	US-10-015-392A-131	Sequence 131, App
C 537	45.2	5.2	21885	3	US-09-949-016-12600	Sequence 12600, A	610	44.6	5.1	2476	5	US-10-011-795B-131	Sequence 131, App
C 538	45.2	5.2	44353	3	US-09-949-016-15302	Sequence 15302, A	611	44.6	5.1	2476	5	US-10-015-386A-131	Sequence 131, App
C 539	45.2	5.2	90923	3	US-09-949-002-623	Sequence 623, App	612	44.6	5.1	2476	5	US-10-012-121A-131	Sequence 131, App
C 540	45.2	5.2	90925	3	US-09-949-002-789	Sequence 789, App	613	44.6	5.1	2476	5	US-10-006-485A-131	Sequence 131, App
C 541	45.2	5.2	103947	3	US-09-949-002-788	Sequence 788, App	614	44.6	5.1	2476	5	US-10-006-746A-131	Sequence 131, App
C 542	45.2	5.2	162841	3	US-09-949-016-13733	Sequence 13733, A	615	44.6	5.1	2476	5	US-10-012-752A-131	Sequence 131, App
C 543	45.2	5.2	188636	3	US-09-949-002-661	Sequence 661, App	616	44.6	5.1	2476	5	US-10-017-253A-131	Sequence 131, App
C 544	45	5.1	555	3	US-09-959-037-2	Sequence 2, Appl1	617	44.6	5.1	2476	5	US-10-015-519A-131	Sequence 131, App
C 545	45	5.1	601	3	US-09-949-016-74713	Sequence 74713, A	618	44.6	5.1	2476	5	US-10-015-715A-131	Sequence 131, App
C 546	45	5.1	601	3	US-09-949-016-74912	Sequence 74912, A	619	44.6	5.1	2476	5	US-10-007-236A-131	Sequence 131, App
C 547	45	5.1	601	3	US-09-949-016-75111	Sequence 75111, A	620	44.6	5.1	141560	3	US-09-949-016-16476	Sequence 16476, A
C 548	45	5.1	601	3	US-09-949-016-75310	Sequence 75310, A	621	44.4	5.1	601	3	US-09-949-016-87162	Sequence 87162, A
C 549	45	5.1	601	3	US-09-949-016-75309	Sequence 75309, A	622	44.4	5.1	662	3	US-08-998-416-185	Sequence 185, App
C 550	45	5.1	601	3	US-09-949-016-106067	Sequence 106067, A	623	44.4	5.1	663	3	US-08-998-416-191	Sequence 191, App
C 551	45	5.1	601	3	US-09-949-016-106266	Sequence 106266, A	624	44.4	5.1	1785	3	US-09-601-198-156	Sequence 156, App
C 552	45	5.1	601	3	US-09-949-016-106465	Sequence 106465, A	625	44.4	5.1	2114	3	US-09-484-970B-165	Sequence 165, App
C 553	45	5.1	601	3	US-09-949-016-106664	Sequence 106664, A	626	44.4	5.1	7015	3	US-09-177-249-6	Sequence 6, Appl1
C 554	45	5.1	601	3	US-09-949-016-106863	Sequence 106863, A	627	44.4	5.1	7015	3	US-09-812-283-6	Sequence 6, Appl1
C 555	45	5.1	1178	3	US-09-270-767-29091	Sequence 29091, A	628	44.4	5.1	7015	5	US-03-071-838A-6	Sequence 6, Appl1
C 556	45	5.1	1186	2	US-08-731-722-5	Sequence 5, Appl1	629	44.4	5.1	85682	4	US-03-531-120-205	Sequence 205, App
C 557	45	5.1	1885	3	US-09-662-254B-41	Sequence 41, Appl1	630	44.4	5.1	101674	3	US-09-949-016-12033	Sequence 12033, A
C 558	45	5.1	2103	3	US-09-270-767-13179	Sequence 13179, A	631	44.4	5.1	126176	3	US-09-949-016-16137	Sequence 16137, A
C 559	44.8	5.1	601	3	US-09-949-016-51986	Sequence 51986, A	632	44.4	5.1	126176	3	US-09-949-016-16138	Sequence 16138, A
C 560	44.8	5.1	711	3	US-08-998-416-786	Sequence 786, App	633	44.4	5.1	258775	3	US-09-949-016-16435	Sequence 16435, A
C 561	44.8	5.1	724	3	US-08-998-416-683	Sequence 683, App	634	44.4	5.1	451924	3	US-09-949-016-12896	Sequence 12896, A
C 562	44.8	5.1	732	3	US-08-998-416-1036	Sequence 1036, App	635	44.4	5.1	451925	3	US-09-949-016-17305	Sequence 17305, A
C 563	44.8	5.1	828	3	US-08-998-416-538	Sequence 538, App	636	44.4	5.1	580073	3	US-08-545-528D-1	Sequence 1, Appl1
C 564	44.8	5.1	828	3	US-08-998-416-538	Sequence 538, App	637	44.4	5.1	611587	3	US-09-531-120-209	Sequence 209, App
C 565	44.8	5.1	1575	3	US-09-543-681A-615	Sequence 615, App	638	44.2	5.0	601	3	US-09-949-016-47890	Sequence 47890, A
C 566	44.8	5.1	2427	3	US-09-601-198-70	Sequence 70, Appl1	639	44.2	5.0	601	3	US-09-949-016-202028	Sequence 202028, A
C 567	44.8	5.1	6336	3	US-09-949-016-12895	Sequence 12895, A	640	44.2	5.0	601	3	US-09-949-016-202030	Sequence 202030, A
C 568	44.8	5.1	6804	3	US-09-949-016-16082	Sequence 16082, A	641	44.2	5.0	700	3	US-09-662-254B-6	Sequence 6, Appl1
C 569	44.8	5.1	6804	3	US-09-949-016-16083	Sequence 16083, A	642	44.2	5.0	879	3	US-09-662-254B-6	Sequence 6, Appl1
C 570	44.8	5.1	21561	3	US-09-949-002-641	Sequence 641, App	643	44.2	5.0	2469	3	US-09-662-254B-32	Sequence 32, Appl1
C 571	44.8	5.1	21561	3	US-09-949-002-797	Sequence 797, App	644	44.2	5.0	3007	3	US-09-710-279-3848	Sequence 3848, App
C 572	44.8	5.1	43676	3	US-09-356-952-12	Sequence 12, Appl1	645	44.2	5.0	3650	3	Sequence 3756, App	Sequence 3756, App
C 573	44.8	5.1	66933	3	US-09-544-398B-11	Sequence 11, Appl1	646	44.2	5.0	4106	3	US-09-710-279-3538	Sequence 3538, App
C 574	44.8	5.1	66933	3	US-09-543-771B-11	Sequence 11, Appl1	647	44.2	5.0	4766	7	PCT-US93-07261-10	Sequence 10, Appl1
C 575	44.8	5.1	72049	3	US-09-544-398B-9	Sequence 9, Appl1	648	44.2	5.0	4920	3	US-08-269-874A-1	Sequence 1, Appl1
C 576	44.8	5.1	72049	3	US-09-543-771B-9	Sequence 9, Appl1	649	44.2	5.0	7454	3	US-08-836-687B-29	Sequence 29, Appl1
C 577	44.8	5.1	129021	4	US-09-531-120-202	Sequence 202, App	650	44.2	5.0	22538	3	US-09-949-016-17405	Sequence 17405, A
C 578	44.8	5.1	129908	3	US-09-585-858-1	Sequence 1, Appl1	651	44.2	5.0	38371	3	US-09-949-016-12061	Sequence 12061, A
C 579	44.8	5.1	129908	3	US-10-270-878-1	Sequence 1, Appl1	652	44.2	5.0	38371	3	US-09-949-016-12488	Sequence 12488, A
C 580	44.8	5.1	14084	3	US-09-949-016-14199	Sequence 14199, A	653	44.2	5.0	38371	3	US-09-949-016-15596	Sequence 15596, A
C 581	44.8	5.1	255679	3	US-09-949-016-17189	Sequence 17189, A	654	44.2	5.0	38371	3	US-09-949-016-15597	Sequence 15597, A
C 582	44.6	5.1	601	3	US-09-949-016-168053	Sequence 168053, A	655	44.2	5.0	137085	3	US-09-949-016-13157	Sequence 13157, A
C 583	44.6	5.1	601	3	US-09-949-016-168054	Sequence 168054, A	656	44.2	5.0	138693	3	US-09-949-016-16724	Sequence 16724, A
C 584	44.6	5.1	601	3	US-09-949-016-168055	Sequence 168055, A	657	44	5.0	2007	3	US-09-248-796A-5247	Sequence 5247, App
C 585	44.6	5.1	601	3	US-09-949-016-168056	Sequence 168056, A	658	44	5.0	3825	4	US-08-704-159-70	Sequence 70, Appl1
C 586	44.6	5.1	1182	3	US-09-601-198-18	Sequence 18, Appl1	659	44	5.0	4441	3	US-08-796-348-20	Sequence 20, Appl1
C 587	44.6	5.1	1316	3	US-09-270-767-2481	Sequence 2481, App	660	44	5.0	6250	3	US-09-724-857-4	Sequence 4, Appl1
C 588	44.6	5.1	1316	3	US-09-270-767-17763	Sequence 17763, App	661	44	5.0	25951	3	US-09-949-016-16194	Sequence 16194, A
C 589	44.6	5.1	1834	3	US-09-270-767-11165	Sequence 11165, A	662	44	5.0	90876	3	US-09-949-016-13271	Sequence 13271, A
C 590	44.6	5.1	1844	3	US-09-123-912-88	Sequence 88, Appl1	663	44	5.0	103935	4	US-09-531-120-193	Sequence 193, App
C 591	44.6	5.1	1844	3	US-09-643-597-88	Sequence 88, App	664	44	5.0	109925	3	US-09-949-016-13210	Sequence 13210, A
C 592	44.6	5.1	1844	3	US-09-480-884A-88	Sequence 88, Appl1	665	44	5.0	137046	3	US-09-949-016-12427	Sequence 12427, A
C 593	44.6	5.1	1844	3	US-09-542-615A-88	Sequence 88, App	666	44	5.0	137046	3	US-09-949-016-13438	Sequence 13438, A
C 594	44.6	5.1	1844	3	US-09-606-421B-88	Sequence 88, Appl1	667	44	5.0	185766	3	US-09-949-002-674	Sequence 674, App
C 595	44.6	5.1	1844	3	US-09-221-107-88	Sequence 88, App	668	44	5.0	185766	3	US-09-949-002-707	Sequence 707, App
C 596	44.6	5.1	1844	3	US-09-466-396A-88	Sequence 88, Appl1	669	44	5.0	212139	3	US-09-949-016-16065	Sequence 16065, A
C 597	44.6	5.1	1844	3	US-09-476-496A-88	Sequence 88, Appl1	670	43.8	5.0	601	3	US-09-949-016-37132	Sequence 37132, A
C 598	44.6	5.1	1844	3	US-09-630-940B-88	Sequence 88, Appl1	671	43.8	5.0	601	3	US-09-949-016-37134	Sequence 37134, A
C 599	44.6	5.1	1844	3	US-09-285-479-88	Sequence 88, Appl1	672	43.8	5.0	601	3	US-09-949-016-47891	Sequence 47891, A
C 600	44.6	5.1	1844	3	US-10-007-700-88	Sequence 88, Appl1	673	43.8	5.0	601	3	US-09-949-016-161109	Sequence 161109, A
C 601	44.6	5.1	2476	3	US-10-012-231A-131	Sequence 131, App	674	43.8	5.0	601	3	US-09-949-016-161111	Sequence 161111, A
C 602	44.6	5.1	2476	3	US-10-015-389A-131	Sequence 131, App	675	43.8	5.0	660	2	US-09-949-016-202029	Sequence 202029, A
C 603	44.6	5.1	2476	3	US-10-006-768A-131	Sequence 131, App	676	43.8	5.0	660	2	US-07-991-867B-32	Sequence 32, Appl1
C 604	44.6	5.1	2476	3	US-10-015-671A-131	Sequence 131, App	677	43.8	5.0	660	2	US-08-107-755A-32	Sequence 32, Appl1
C 605	44.6	5.1	2476	3	US-10-015-393A-131	Sequence 131, App	678	43.8	5.0	660	2	US-08-544-332-32	Sequence 32, Appl1
C 606	44.6	5.1	2476	3	US-10-011-833A-131	Sequence 131, App	679	43.8	5.0	660	3	US-09-370-861A-32	Sequence 32, Appl1
C 607	44.6	5.1	2476	3	US-10-006-041A-131	Sequence 131, App	680	43.8	5.0	2093	3	US-09-269-137-7	Sequence 7, Appl1

681	43.8	5.0	2619	3	US-09-662-254B-46	Sequence 46, Appl	C 754	43.2	4.9	3387	2	US-08-261-822A-5	Sequence 5, Appl
682	43.8	5.0	4810	3	US-08-852-629-11	Sequence 11, Appl	C 755	43.2	4.9	3387	7	PCT-US95-07744A-5	Sequence 5, Appl
683	43.8	5.0	4838	3	US-08-852-629-15	Sequence 15, Appl	C 756	43.2	4.9	17082	3	US-09-949-016-14893	Sequence 14893, A
c 684	43.8	5.0	5006	4	US-09-837-751-7	Sequence 7, Appl	C 757	43.2	4.9	17865	3	US-09-949-016-15950	Sequence 15950, A
685	43.8	5.0	57875	3	US-09-949-016-13152	Sequence 13152, A	C 758	43.2	4.9	30246	3	US-08-956-171B-56	Sequence 56, Appl
686	43.8	5.0	101674	3	US-09-949-016-12033	Sequence 12033, A	C 759	43.2	4.9	30246	3	US-08-781-986A-56	Sequence 56, Appl
c 687	43.8	5.0	104428	3	US-09-949-016-12737	Sequence 12737, A	C 760	43.2	4.9	32669	3	US-09-949-016-17250	Sequence 17250, A
c 688	43.8	5.0	104429	3	US-09-949-016-13814	Sequence 13814, A	C 761	43.2	4.9	64638	3	US-09-949-016-11767	Sequence 11767, A
689	43.8	5.0	105001	3	US-09-949-002-627	Sequence 627, App	C 762	43.2	4.9	64639	3	US-09-949-016-13520	Sequence 13520, A
690	43.8	5.0	105002	3	US-09-949-002-821	Sequence 821, App	C 763	43.2	4.9	601	3	US-09-949-016-13520	Sequence 13520, A
c 691	43.8	5.0	151828	4	US-09-531-120-197	Sequence 197, App	C 764	43.2	4.9	601	3	US-09-949-016-46909	Sequence 46909, A
c 692	43.8	5.0	153866	3	US-09-949-016-16919	Sequence 16919, A	C 765	43.2	4.9	601	3	US-09-949-016-157756	Sequence 157756, A
c 693	43.8	5.0	187169	3	US-09-949-016-12776	Sequence 12776, A	C 766	43.2	4.9	601	3	US-09-949-016-179399	Sequence 179399, A
c 694	43.8	5.0	191569	3	US-09-949-016-15940	Sequence 15940, A	C 767	43.2	4.9	951	4	US-09-747-385-14	Sequence 14, Appl
695	43.6	5.0	456	3	US-09-328-352-2410	Sequence 2410, Ap	C 768	43.2	4.9	1341	3	US-09-830-230A-572	Sequence 572, App
696	43.6	5.0	601	3	US-09-949-016-51985	Sequence 51985, A	C 769	43.2	4.9	1410	3	US-09-830-230A-571	Sequence 571, App
697	43.6	5.0	601	3	US-09-949-016-51987	Sequence 51987, A	C 770	43.2	4.9	1864	3	US-09-468-265-4	Sequence 4, Appl
698	43.6	5.0	601	3	US-09-949-016-60639	Sequence 60639, A	C 771	43.2	4.9	2863	3	US-09-270-767-13299	Sequence 13299, A
699	43.6	5.0	601	3	US-09-949-016-171661	Sequence 171661, A	C 772	43.2	4.9	2447	2	US-09-014-969-14	Sequence 14, Appl
700	43.6	5.0	601	3	US-09-949-016-171662	Sequence 171662, A	C 773	43.2	4.9	2861	2	US-08-299-953-1	Sequence 1, Appl
701	43.6	5.0	1128	3	US-09-134-003C-2242	Sequence 2242, Ap	C 774	43.2	4.9	2861	3	US-08-459-415-1	Sequence 1, Appl
c 702	43.6	5.0	1374	3	US-09-601-198-158	Sequence 158, App	C 775	43.2	4.9	2861	3	US-09-066-687-1	Sequence 1, Appl
703	43.6	5.0	1410	3	US-09-662-254B-33	Sequence 33, Appl	C 776	43.2	4.9	2861	7	PCT-US95-11231-1	Sequence 1, Appl
704	43.6	5.0	1738	3	US-09-918-909A-27	Sequence 27, Appl	C 777	43.2	4.9	3607	2	US-08-647-351B-1	Sequence 1, Appl
705	43.6	5.0	2453	3	US-09-710-279-4379	Sequence 4379, Ap	C 778	43.2	4.9	3607	3	US-09-326-074-1	Sequence 1, Appl
c 706	43.6	5.0	2853	3	US-09-509-188-3	Sequence 3, Appl	C 779	43.2	4.9	3834	3	US-09-209-668-18	Sequence 18, Appl
c 707	43.6	5.0	7970	3	US-09-949-016-13277	Sequence 13277, A	C 780	43.2	4.9	3834	3	US-09-949-016-71	Sequence 71, Appl
708	43.6	5.0	12482	3	US-09-512-563C-25	Sequence 25, Appl	C 781	43.2	4.9	3834	5	US-09-943-679A-2479	Sequence 2479, Ap
709	43.6	5.0	12482	3	US-10-060-607-25	Sequence 25, Appl	C 782	43.2	4.9	3854	2	US-08-365-470-1	Sequence 1, Appl
c 710	43.6	5.0	38371	3	US-09-949-016-12061	Sequence 12061, A	C 783	43.2	4.9	3858	2	US-08-344-155C-98	Sequence 98, Appl
c 711	43.6	5.0	38371	3	US-09-949-016-12488	Sequence 12488, A	C 784	43.2	4.9	3858	3	US-09-009-490A-88	Sequence 88, Appl
c 712	43.6	5.0	38371	3	US-09-949-016-15596	Sequence 15596, A	C 785	43.2	4.9	3858	3	US-09-982-262C-89	Sequence 89, Appl
c 713	43.6	5.0	38371	3	US-09-949-016-15597	Sequence 15597, A	C 786	43.2	4.9	3862	3	US-09-949-016-4401	Sequence 4401, Ap
714	43.6	5.0	50073	3	US-09-949-016-16026	Sequence 16026, A	C 787	43.2	4.9	3863	3	US-08-482-073-1	Sequence 1, Appl
715	43.6	5.0	133278	3	US-09-949-016-12524	Sequence 12524, A	C 788	43.2	4.9	3863	10	5217870-1	Patent No. 5217870
c 716	43.6	5.0	181429	3	US-09-949-016-12372	Sequence 12372, A	C 789	43.2	4.9	3881	2	US-08-299-953-2	Sequence 2, Appl
c 717	43.6	5.0	181430	3	US-09-949-016-15772	Sequence 15772, A	C 790	43.2	4.9	3881	2	US-08-459-415-2	Sequence 2, Appl
c 718	43.6	5.0	225127	3	US-09-949-016-16480	Sequence 16480, A	C 791	43.2	4.9	3881	3	US-09-066-687-2	Sequence 2, Appl
719	43.6	5.0	283538	3	US-09-949-016-13506	Sequence 13506, A	C 792	43.2	4.9	3881	7	PCT-US95-11231-2	Sequence 2, Appl
720	43.6	5.0	331814	3	US-09-949-016-12008	Sequence 12008, A	C 793	43.2	4.9	4167	3	US-09-973-278-700	Sequence 700, App
721	43.6	5.0	331814	3	US-09-949-016-17056	Sequence 17056, A	C 794	43.2	4.9	15384	3	US-09-949-016-11813	Sequence 11813, A
722	43.4	5.0	293	3	US-09-270-767-9068	Sequence 9068, Ap	C 795	43.2	4.9	15413	3	US-09-949-016-16143	Sequence 16143, A
723	43.4	5.0	591	3	US-09-248-796A-7131	Sequence 24350, A	C 796	43.2	4.9	16438	3	US-09-949-016-16165	Sequence 16165, A
c 724	43.4	5.0	601	3	US-09-949-016-46905	Sequence 46905, A	C 797	43.2	4.9	18989	3	US-09-949-016-13435	Sequence 13435, A
c 725	43.4	5.0	601	3	US-09-949-016-46906	Sequence 46906, A	C 798	43.2	4.9	19698	3	US-09-949-016-14079	Sequence 14079, A
c 726	43.4	5.0	601	3	US-09-949-016-46907	Sequence 46907, A	C 799	43.2	4.9	21885	3	US-09-949-016-12600	Sequence 12600, A
c 727	43.4	5.0	601	3	US-09-949-016-46908	Sequence 46908, A	C 800	43.2	4.9	42000	3	US-10-081-563-25	Sequence 25, Appl
c 728	43.4	5.0	601	3	US-09-949-016-47329	Sequence 47329, A	C 801	43.2	4.9	50381	3	US-09-949-016-17122	Sequence 17122, A
729	43.4	5.0	601	3	US-09-949-016-47329	Sequence 47329, A	C 802	43.2	4.9	58407	3	US-08-916-421B-2	Sequence 2, Appl
730	43.4	5.0	601	3	US-09-949-016-47734	Sequence 47734, A	C 803	43.2	4.9	58407	3	US-09-592-570-2	Sequence 2, Appl
731	43.4	5.0	601	3	US-09-949-016-166398	Sequence 166398, A	C 804	43.2	4.9	69909	3	US-09-949-016-13423	Sequence 13423, A
732	43.4	5.0	601	3	US-09-949-016-179398	Sequence 179398, A	C 805	43.2	4.9	107800	3	US-09-949-016-13118	Sequence 13118, A
733	43.4	5.0	834	3	US-08-998-416-305	Sequence 305, App	C 806	43.2	4.9	111235	3	US-09-949-016-15328	Sequence 15328, A
734	43.4	5.0	900	3	US-09-707-767-11765	Sequence 11765, A	C 807	43.2	4.9	116425	3	US-09-949-016-11809	Sequence 11809, A
735	43.4	5.0	1032	3	US-09-270-767-14663	Sequence 14663, A	C 808	43.2	4.9	141589	5	US-09-543-679A-2480	Sequence 2480, Ap
c 736	43.4	5.0	1865	3	US-09-601-198-104	Sequence 104, App	C 809	43.2	4.9	141589	5	US-09-543-679A-2694	Sequence 2694, Ap
c 737	43.4	5.0	11375	3	US-09-949-016-16047	Sequence 16047, A	C 810	43.2	4.9	141589	5	US-09-543-679A-2719	Sequence 2719, Ap
738	43.4	5.0	19557	7	PCT-US92-06300-1	Sequence 1, Appl	C 811	43.2	4.9	146982	5	US-09-543-679A-3009	Sequence 3009, Ap
c 739	43.4	5.0	113966	3	US-09-949-016-12277	Sequence 12277, A	C 812	43.2	4.9	146982	5	US-09-543-679A-3009	Sequence 3009, Ap
c 740	43.4	5.0	113967	3	US-09-949-016-17051	Sequence 17051, A	C 813	43.2	4.9	209274	5	US-09-543-679A-3004	Sequence 3004, Ap
c 741	43.4	5.0	128175	3	US-09-949-016-16268	Sequence 16268, A	C 814	43.2	4.9	392000	3	US-10-027-983-11	Sequence 11, Appl
742	43.4	5.0	167708	3	US-09-949-016-16423	Sequence 16423, A	C 815	42.8	4.9	465	3	US-09-270-767-11456	Sequence 11456, A
c 743	43.4	5.0	246444	3	US-09-949-016-13113	Sequence 13113, A	C 816	42.8	4.9	900	3	US-09-543-681A-1453	Sequence 1453, Ap
c 744	43.2	4.9	601	3	US-09-949-016-186364	Sequence 186364, A	C 817	42.8	4.9	1117	3	US-09-247-373B-33	Sequence 33, Appl
c 745	43.2	4.9	601	3	US-09-949-016-193741	Sequence 193741, A	C 818	42.8	4.9	1368	3	US-09-248-796A-11323	Sequence 11323, A
c 746	43.2	4.9	731	2	US-08-451-405A-2	Sequence 2, Appl	C 819	42.8	4.9	1866	3	US-10-015-637-13	Sequence 13, Appl
c 747	43.2	4.9	767	3	US-08-998-416-472	Sequence 472, App	C 820	42.8	4.9	2118	3	US-09-601-198-59	Sequence 59, Appl
748	43.2	4.9	972	3	US-09-286-690-1	Sequence 1, Appl	C 821	42.8	4.9	3047	3	US-09-710-279-4113	Sequence 4113, Ap
749	43.2	4.9	1269	3	US-09-601-198-178	Sequence 178, App	C 822	42.8	4.9	3263	3	US-08-714-918-20	Sequence 20, Appl
c 750	43.2	4.9	1353	3	US-09-601-198-37	Sequence 37, Appl	C 823	42.8	4.9	3263	3	US-09-265-315-20	Sequence 20, Appl
c 751	43.2	4.9	1373	3	US-09-270-767-5756	Sequence 5756, Ap	C 824	42.8	4.9	3263	3	US-09-265-315-20	Sequence 20, Appl
752	43.2	4.9	1373	3	US-09-270-767-21038	Sequence 21038, A	C 825	42.8	4.9	3263	3	US-09-266-417-20	Sequence 20, Appl
753	43.2	4.9	2469	3	US-09-248-796A-4746	Sequence 4746, Ap	C 826	42.8	4.9	3263	3	US-09-528-709-20	Sequence 20, Appl

c 827	42.8	4.9	3263	3	US-09-527-745-20	Sequence 20, Appl	900	42.4	4.8	2216	7	PCT-US94-12364-24	Sequence 24, Appl
c 828	42.8	4.9	3450	3	US-09-562-254B-47	Sequence 47, Appl	901	42.4	4.8	2216	7	PCT-US95-07753-3	Sequence 3, Appl
c 829	42.8	4.9	6768	4	US-08-107-755A-1	Sequence 1, Appl	902	42.4	4.8	2230	3	US-08-378-313-24	Sequence 24, Appl
c 830	42.8	4.9	8457	2	US-07-991-867B-1	Sequence 1, Appl	903	42.4	4.8	2298	3	US-09-157-077-7	Sequence 7, Appl
c 831	42.8	4.9	8457	2	US-08-544-332-1	Sequence 1, Appl	904	42.4	4.8	2327	3	US-09-157-077-1	Sequence 1, Appl
c 832	42.8	4.9	8457	2	US-09-370-861A-1	Sequence 1, Appl	905	42.4	4.8	3131	3	US-09-710-279-4221	Sequence 4221, Ap
c 833	42.8	4.9	10396	2	US-08-245-809-5	Sequence 5, Appl	906	42.4	4.8	3258	3	US-09-830-230A-324	Sequence 324, App
c 834	42.8	4.9	10965	2	US-08-107-748-4	Sequence 4, Appl	c 907	42.4	4.8	3272	3	US-09-796-348-18	Sequence 18, Appl
c 835	42.8	4.9	10965	7	PCT-US92-01385-4	Sequence 4, Appl	c 908	42.4	4.8	3303	3	US-09-796-348-19	Sequence 19, Appl
c 836	42.8	4.9	22786	3	US-09-949-016-13399	Sequence 13399, A	c 909	42.4	4.8	3315	3	US-09-796-348-17	Sequence 17, Appl
c 837	42.8	4.9	22846	2	US-08-469-461-3	Sequence 3, Appl	c 910	42.4	4.8	3354	3	US-09-830-230A-323	Sequence 323, App
c 838	42.8	4.9	22846	3	US-07-890-609-3	Sequence 3, Appl	c 911	42.4	4.8	4441	3	US-09-796-348-20	Sequence 20, Appl
c 839	42.8	4.9	24953	3	US-09-949-016-15743	Sequence 15743, A	c 912	42.4	4.8	4673	2	US-07-638-431-1	Sequence 1, Appl
c 840	42.8	4.9	34011	3	US-09-949-016-12485	Sequence 12485, A	c 913	42.4	4.8	4673	7	PCT-US92-00018-1	Sequence 11, Appl
c 841	42.8	4.9	42250	3	US-09-949-016-17171	Sequence 17171, A	c 914	42.4	4.8	4810	3	US-08-852-629-11	Sequence 11, Appl
c 842	42.8	4.9	51698	3	US-09-949-016-12671	Sequence 12671, A	c 915	42.4	4.8	4838	3	US-08-852-629-15	Sequence 15, Appl
c 843	42.8	4.9	87080	3	US-09-949-016-17011	Sequence 17011, A	c 916	42.4	4.8	5887	4	US-09-747-385-6	Sequence 6, Appl
c 844	42.8	4.9	90428	3	US-09-949-016-12564	Sequence 12564, A	c 917	42.4	4.8	8700	2	US-08-392-625-16	Sequence 16, Appl
c 845	42.8	4.9	91665	3	US-09-949-016-12234	Sequence 12234, A	c 918	42.4	4.8	8700	2	US-08-466-961A-16	Sequence 16, Appl
c 846	42.8	4.9	113538	3	US-09-949-016-16329	Sequence 16329, A	c 919	42.4	4.8	8700	2	US-08-645-193B-18	Sequence 18, Appl
c 847	42.8	4.9	134140	3	US-09-949-016-12672	Sequence 12672, A	c 920	42.4	4.8	15418	3	US-09-783-203-1	Sequence 1, Appl
c 848	42.8	4.9	134241	3	US-09-949-016-12824	Sequence 12824, A	c 921	42.4	4.8	15418	3	US-09-934-427A-1	Sequence 1, Appl
c 849	42.8	4.9	134242	3	US-09-949-016-15813	Sequence 15813, A	c 922	42.4	4.8	15418	3	US-09-244-438-1	Sequence 1, Appl
c 850	42.8	4.9	134242	3	US-09-949-016-15814	Sequence 15814, A	c 923	42.4	4.8	16573	3	US-09-995-419A-1	Sequence 1, Appl
c 851	42.8	4.9	134242	3	US-09-949-016-15815	Sequence 15815, A	c 924	42.4	4.8	16573	3	US-09-949-016-14876	Sequence 14876, A
c 852	42.8	4.9	163317	4	US-09-531-120-212	Sequence 212, App	c 925	42.4	4.8	60589	3	US-09-949-016-17070	Sequence 17070, A
c 853	42.8	4.9	174170	3	US-09-949-016-14810	Sequence 14810, A	c 926	42.4	4.8	98844	3	US-09-791-211-10	Sequence 10, Appl
c 854	42.8	4.9	174170	3	US-09-949-016-14811	Sequence 14811, A	c 927	42.4	4.8	101894	3	US-09-949-016-12005	Sequence 12005, A
c 855	42.8	4.9	174318	3	US-09-949-016-11880	Sequence 11880, A	c 928	42.4	4.8	103894	3	US-09-949-016-14450	Sequence 14450, A
c 856	42.8	4.9	174318	3	US-09-949-016-14812	Sequence 14812, A	c 929	42.4	4.8	112465	3	US-09-949-016-14450	Sequence 14450, A
c 857	42.8	4.9	174318	3	US-09-949-016-14813	Sequence 14813, A	c 930	42.4	4.8	134890	3	US-09-949-016-15602	Sequence 15602, A
c 858	42.6	4.9	601	3	US-09-949-016-31621	Sequence 31621, A	c 931	42.4	4.8	143776	3	US-09-949-001-29	Sequence 29, Appl
c 859	42.6	4.9	601	3	US-09-949-016-37133	Sequence 37133, A	c 932	42.4	4.8	144034	3	US-09-949-001-35	Sequence 35, Appl
c 860	42.6	4.9	601	3	US-09-949-016-37135	Sequence 37135, A	c 933	42.4	4.8	157032	3	US-09-949-016-16502	Sequence 16502, A
c 861	42.6	4.9	601	3	US-09-949-016-146639	Sequence 146639, A	c 934	42.4	4.8	167677	3	US-09-949-016-12147	Sequence 12147, A
c 862	42.6	4.9	601	3	US-09-949-016-150427	Sequence 150427, A	c 935	42.4	4.8	167677	3	US-09-949-016-17361	Sequence 17361, A
c 863	42.6	4.9	601	3	US-09-949-016-161110	Sequence 161110, A	c 936	42.4	4.8	601	3	US-09-949-016-30532	Sequence 30532, A
c 864	42.6	4.9	601	3	US-09-949-016-161112	Sequence 161112, A	c 937	42.2	4.8	601	3	US-09-949-016-37151	Sequence 37151, A
c 865	42.6	4.9	945	3	US-09-601-198-177	Sequence 177, App	c 938	42.2	4.8	601	3	US-09-949-016-37165	Sequence 37165, A
c 866	42.6	4.9	984	3	US-09-543-681A-4159	Sequence 4159, App	c 939	42.2	4.8	601	3	US-09-949-016-47330	Sequence 47330, A
c 867	42.6	4.9	1692	3	US-09-601-198-63	Sequence 63, Appl	c 940	42.2	4.8	601	3	US-09-949-016-47735	Sequence 47735, A
c 868	42.6	4.9	2475	3	US-09-220-132-179	Sequence 179, App	c 941	42.2	4.8	601	3	US-09-949-016-92234	Sequence 92234, A
c 869	42.6	4.9	4152	3	US-09-562-254B-9	Sequence 9, Appl	c 942	42.2	4.8	601	3	US-09-949-016-145869	Sequence 145869, A
c 870	42.6	4.9	4185	3	US-09-417-485D-7	Sequence 7, Appl	c 943	42.2	4.8	601	3	US-09-949-016-146137	Sequence 146137, A
c 871	42.6	4.9	4270	3	US-09-949-016-12549	Sequence 12549, A	c 944	42.2	4.8	601	3	US-09-949-016-146405	Sequence 146405, A
c 872	42.6	4.9	4336	3	US-09-949-016-13959	Sequence 13959, A	c 945	42.2	4.8	601	3	US-09-949-016-179398	Sequence 179398, A
c 873	42.6	4.9	4818	3	US-08-817-926-27	Sequence 27, Appl	c 946	42.2	4.8	609	3	US-09-248-796A-1998	Sequence 1998, Ap
c 874	42.6	4.9	17612	3	US-09-949-016-15061	Sequence 15061, A	c 947	42.2	4.8	613	2	US-08-740-309-1	Sequence 1, Appl
c 875	42.6	4.9	46823	3	US-09-949-016-12723	Sequence 12723, A	c 948	42.2	4.8	714	3	US-09-662-254B-51	Sequence 51, Appl
c 876	42.6	4.9	46940	3	US-09-949-016-16252	Sequence 16252, A	c 949	42.2	4.8	831	3	US-09-662-254B-69	Sequence 69, Appl
c 877	42.6	4.9	139049	3	US-09-949-016-17030	Sequence 17030, A	c 950	42.2	4.8	1230	4	US-09-747-385-5	Sequence 5, Appl
c 878	42.6	4.9	373182	3	US-09-949-016-17371	Sequence 17371, A	c 951	42.2	4.8	1314	3	US-09-248-796A-8070	Sequence 8070, Ap
c 879	42.6	4.9	373694	3	US-09-949-016-12062	Sequence 12062, A	c 952	42.2	4.8	1473	3	US-09-830-230A-440	Sequence 440, App
c 880	42.6	4.9	580073	3	US-08-545-528D-1	Sequence 1, Appl	c 953	42.2	4.8	1515	3	US-09-248-796A-9828	Sequence 9828, Ap
c 881	42.4	4.8	225	3	US-09-543-681A-1045	Sequence 1045, Ap	c 954	42.2	4.8	1526	4	US-08-704-159-43	Sequence 43, Appl
c 882	42.4	4.8	601	3	US-09-949-016-37136	Sequence 37136, A	c 955	42.2	4.8	1527	3	US-09-830-230A-439	Sequence 439, App
c 883	42.4	4.8	601	3	US-09-949-016-161113	Sequence 161113, A	c 956	42.2	4.8	1864	3	US-09-468-265-4	Sequence 4, Appl
c 884	42.4	4.8	601	3	US-09-949-016-185650	Sequence 185650, A	c 957	42.2	4.8	2469	3	US-09-662-254B-32	Sequence 32, Appl
c 885	42.4	4.8	601	3	US-09-949-016-186528	Sequence 186528, A	c 958	42.2	4.8	3318	3	US-09-662-254B-7	Sequence 7, Appl
c 886	42.4	4.8	660	2	US-07-991-867B-32	Sequence 32, Appl	c 959	42.2	4.8	3701	3	US-08-845-258-10	Sequence 10, Appl
c 887	42.4	4.8	660	2	US-08-107-755A-32	Sequence 32, Appl	c 960	42.2	4.8	3701	3	US-08-990-571-10	Sequence 10, Appl
c 888	42.4	4.8	660	2	US-08-544-332-32	Sequence 32, Appl	c 961	42.2	4.8	3701	3	US-08-723-142A-10	Sequence 10, Appl
c 889	42.4	4.8	660	3	US-09-370-861A-32	Sequence 32, Appl	c 962	42.2	4.8	3701	3	US-09-528-784A-10	Sequence 10, Appl
c 890	42.4	4.8	854	3	US-08-998-416-534	Sequence 534, App	c 963	42.2	4.8	3701	3	US-09-569-098A-10	Sequence 10, Appl
c 891	42.4	4.8	1603	3	US-09-009-443-11	Sequence 11, Appl	c 964	42.2	4.8	3876	4	US-08-704-159-39	Sequence 39, Appl
c 892	42.4	4.8	1785	3	US-09-117-415B-1	Sequence 1, Appl	c 965	42.2	4.8	6996	5	US-09-573-080A-29	Sequence 29, Appl
c 893	42.4	4.8	1842	3	US-09-117-415B-21	Sequence 21, Appl	c 966	42.2	4.8	6996	5	US-09-854-867-29	Sequence 29, Appl
c 894	42.4	4.8	1896	3	US-09-117-415B-17	Sequence 17, Appl	c 967	42.2	4.8	29628	3	US-09-949-016-16483	Sequence 16483, A
c 895	42.4	4.8	1896	3	US-09-117-415B-19	Sequence 19, Appl	c 968	42.2	4.8	58829	3	US-09-949-016-13146	Sequence 13146, A
c 896	42.4	4.8	1950	3	US-09-117-415B-15	Sequence 15, Appl	c 969	42.2	4.8	59519	3	US-09-949-016-13504	Sequence 13504, A
c 897	42.4	4.8	2216	2	US-08-046-583-12	Sequence 12, Appl	c 970	42.2	4.8	114426	3	US-09-949-016-15078	Sequence 15078, A
c 898	42.4	4.8	2216	2	US-08-384-556A-3	Sequence 3, Appl	c 971	42.2	4.8	134499	4	US-09-531-120-192	Sequence 192, App
c 899	42.4	4.8	2216	2	US-08-331-355A-24	Sequence 24, Appl	c 972	42.2	4.8	143248	3	US-09-949-016-16652	Sequence 16652, A

```
c 973 42.2 4.8 186959 3 US-09-949-016-13125 Sequence 13125, A
974 42 4.8 601 3 US-09-949-016-18854 Sequence 18854, A
975 42 4.8 601 3 US-09-949-016-56719 Sequence 56719, A
c 976 42 4.8 601 3 US-09-949-016-148052 Sequence 148052, A
977 42 4.8 601 3 US-09-949-016-148158 Sequence 148158, A
978 42 4.8 601 3 US-09-949-016-168053 Sequence 168053, A
979 42 4.8 601 3 US-09-949-016-168054 Sequence 168054, A
980 42 4.8 601 3 US-09-949-016-168055 Sequence 168055, A
981 42 4.8 601 3 US-09-949-016-188056 Sequence 188056, A
c 982 42 4.8 601 3 US-09-949-016-196403 Sequence 196403, A
983 42 4.8 601 3 US-09-949-016-198782 Sequence 198782, A
984 42 4.8 1522 3 US-09-056-783-1 Sequence 1, Appli
c 985 42 4.8 3078 3 US-09-601-198-54 Sequence 54, Appli
986 42 4.8 3138 2 US-07-867-106-4 Sequence 4, Appli
987 42 4.8 3906 3 US-09-662-254B-60 Sequence 60, Appli
988 42 4.8 5750 3 US-09-949-016-14714 Sequence 14714, A
989 42 4.8 5750 3 US-09-949-016-14715 Sequence 14715, A
c 990 42 4.8 17612 3 US-09-949-016-15061 Sequence 15061, A
c 991 42 4.8 21723 3 US-09-949-016-16383 Sequence 16383, A
c 992 42 4.8 23501 3 US-09-949-016-17517 Sequence 17517, A
c 993 42 4.8 26000 3 US-09-843-378-10 Sequence 10, Appli
c 994 42 4.8 47363 3 US-09-949-016-13420 Sequence 13420, A
c 995 42 4.8 93493 3 US-09-949-016-12063 Sequence 12063, A
c 996 42 4.8 94156 3 US-09-949-016-12388 Sequence 12388, A
997 42 4.8 94618 4 US-09-531-120-191 Sequence 191, App
c 998 42 4.8 95020 3 US-09-949-016-13272 Sequence 13272, A
c 999 42 4.8 102406 3 US-09-949-016-14673 Sequence 14673, A
c1000 42 4.8 141560 3 US-09-949-016-16476 Sequence 16476, A
```

## ALIGNMENTS

```
RESULT 1
US-09-495-406-8
; Sequence 8, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni O:19
US-09-495-406-8
```

```
Query Match 100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 5.7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATTCATTATCAAGG 60
DB 1 ATGAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATTCATTATCAAGG 60
QY 61 CTACCAAAATGATTTCATGATGTTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
DB 61 CTACCAAAATGATTTCATGATGTTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
```

```
QY 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
DB 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 181 ACTTTAAAAACATTTAATCCAAATATGAGACCGAACTAAATATGTTGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAATCCAAATATGAGACCGAACTAAATATGTTGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACAGATTATTTTCCT 300
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACAGATTATTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAATTTAATGCTTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAATTTAATGCTTTATTTAAA 360
QY 361 TTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
DB 361 TTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAAT 540
DB 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTACAAATATAAATATATGCTTATGCTTAATAGTCTTTTACGAAATTTT 660
DB 601 GAAAAAATTTACAAATATAAATATATGCTTATGCTTAATAGTCTTTTACGAAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACT 720
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780
DB 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780
QY 781 AAAAAATATAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATT 840
DB 781 AAAAAATATAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATT 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876
DB 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876
```

## RESULT 2

```
US-09-816-028A-8
; Sequence 8, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
```

```
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-816-028A-8

Query Match      100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 5.7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180

Qy 181 ACTTTAAAAATTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240
Db 181 ACTTTAAAAATTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAAGAAATTTAAAGAAATTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAAGAAATTTAAAGAAATTTTAA 360

Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTAAACAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTAAACAAATGGGTCA 480

Qy 481 TCCTATGCTTTTGATACCAAGAAATTTCTTTTAAAAATCTAGCCCTGATTTTAAAAAT 540
Db 481 TCCTATGCTTTTGATACCAAGAAATTTCTTTTAAAAATCTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 601 GAAAAAATTTTAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Db 601 GAAAAAATTTTAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660

Qy 661 ATAGAACTAGCGCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 720
Db 661 ATAGAACTAGCGCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 720

Qy 721 AAAGATATCTCATACCTTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTCATACCTTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTTT 780

Qy 781 AAAAAAATAAATAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAGATTA 840
Db 781 AAAAAAATAAATAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAGATTA 840

Qy 841 CCTAGTATTAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTATTAAGCAATTTTCAAGGAAATAA 876
```

```
RESULT 3
US-10-303-162-8
; Sequence 8, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: Gangliosides and Ganglioside Mimics
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-162-8
```

```
Query Match      100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 5.7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180

Qy 181 ACTTTAAAAATTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240
Db 181 ACTTTAAAAATTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAAGAAATTTAAAGAAATTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAAGAAATTTAAAGAAATTTTAA 360

Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTAAACAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTAAACAAATGGGTCA 480

Qy 481 TCCTATGCTTTTGATACCAAGAAATTTCTTTTAAAAATCTAGCCCTGATTTTAAAAAT 540
Db 481 TCCTATGCTTTTGATACCAAGAAATTTCTTTTAAAAATCTAGCCCTGATTTTAAAAAT 540
```

```
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAAATCTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAAATCTACACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTTCAAAAATATTAATTTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTTCAAAAATATTAATTTT 780
QY 781 AAAAAAATAAAAAATTAAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAGATT 840
Db 781 AAAAAAATAAAAAATTAAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAGATT 840
QY 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAATAA 876

RESULT 4
US-10-303-134-8
; Sequence 8, Application US/103031134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-134-8

Query Match 100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred No. 5,7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATGATTATTCGAAG 60
Db 1 ATGAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATGATTATTCGAAG 60
QY 61 CTACCAATGATTTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTTGGTAAAAAATGCAAAACAGTGTTTACACCCCTAAATTCCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTACACCCCTAAATTCCTTTGAGCAATACTAC 180
```

```
QY 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTAATGCTTATTTTAAA 360
QY 361 TTTTACGAAATTTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAGTAGCC 420
Db 361 TTTTACGAAATTTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTCATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAAATTTCTA 600
QY 601 GAAAAAATTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAAATCTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAAATCTACACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTTCAAAAATATTAATTTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTTCAAAAATATTAATTTT 780
QY 781 AAAAAAATAAAAAATTAAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAGATT 840
Db 781 AAAAAAATAAAAAATTAAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAGATT 840
QY 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAATAA 876
```

```
RESULT 5
US-10-303-118-8
; Sequence 8, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
```



```
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; -10-303-118-8
US-10-303-118-8

Query Match      100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 5.7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTAAAGAAATTTAAAGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTAAAGAAATTTAAAGCTTATTTTAAA 360

Qy 361 TTTTACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Qy 421 ATAGCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAAGAAATTTTATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAGAAATTTTATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

Qy 541 GATGCTCGGACATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATGCTCGGACATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTATGCTTATGCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTATGCTTATGCTTTTACGAAATTTT 660

Qy 720 ATAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATCAAGAAATAAATACTACACT 720
Db 720 ATAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATCAAGAAATAAATACTACACT 720

Qy 780 AAAGATATACTCATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780
Db 780 AAAGATATACTCATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780

Qy 840 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAGTTGATAAAGATCTTATTAAGATTA 840
Db 840 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAGTTGATAAAGATCTTATTAAGATTA 840

Qy 841 CCTAGTGATATAAGCATTTTCAAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCATTTTCAAAGGAAATAA 876
```

```
RESULT 6
US-10-303-128-8
; Sequence 8, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-006111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-128-8
```

```
Query Match      100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 5.7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTAAAGAAATTTAAAGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTAAAGAAATTTAAAGCTTATTTTAAA 360

Qy 361 TTTTACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Qy 421 ATAGCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAAGAAATTTTATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAGAAATTTTATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
```





```
; NAME/KEY: CDS
; LOCATION: (1)...(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-09-816-028A-13

Query Match          99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG 60

Qy 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAACTAT 120
Db 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAACTAT 120

Qy 121 CTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180
Db 121 CTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTTTAAAGAAATTTTAAAGAAAT 360
Db 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTTTAAAGAAATTTTAAAGAAAT 360

Qy 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTAAAGAAATTTTAAAG 420
Db 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTAAAGAAATTTTAAAG 420

Qy 421 ATAGCACTAGCGCAATTTAAATTTTAAATTTTATCATACAGAAAAAATACTACACT 720
Db 421 ATAGCACTAGCGCAATTTAAATTTTAAATTTTATCATACAGAAAAAATACTACACT 720

Qy 721 AAAGATATCTCATACCTCTAGTGAGCTTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTCATACCTCTAGTGAGCTTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780

Qy 781 AAAAAATAAAAAATTAAGAAAAATGTTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAAATTAAGAAAAATGTTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCAATTTATTTCAAGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTATTTCAAGAAAA 873
```

RESULT 9  
US-10-303-162-13

```
; Sequence 13, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-162-13
```

```
Query Match          99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG 60

Qy 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAACTAT 120
Db 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAACTAT 120

Qy 121 CTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180
Db 121 CTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTTTAAAGAAATTTTAAAGAAAT 360
Db 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTTTAAAGAAATTTTAAAGAAAT 360

Qy 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTAAAGAAATTTTAAAG 420
Db 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTAAAGAAATTTTAAAG 420

Qy 421 ATAGCCCTAGGATCAAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATCAAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
Db 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540

Qy 541 GATCGCTGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
```

```
QY 601 GAAAAAATTACAAAATAAACTATATGCTTATGCTCCTAAATAGTCTCTTTAGCAAAATTTT 660
Db 601 GAAAAAATTACAAAATAAACTATATGCTTATGCTCCTAAATAGTCTCTTTAGCAAAATTTT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAGAAAAAATAACTACACT 720

QY 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780

QY 781 AAAAAAATAAATAAATAAAGAAATGTTTATACAAGTTGATATAAGACATCTATTAAAGATTA 840
Db 781 AAAAAAATAAATAAATAAAGAAATGTTTATACAAGTTGATATAAGACATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873

RESULT 10
US-10-303-134-13
; Sequence 13, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CatII) from C. jejuni O:36
US-10-303-134-13

Query Match          99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60

QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTGTGTAATAAATGCAACAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAACAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTTAAATCCAAAATCAAGAATATGAGACCGAACAATAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAAATCAAGAATATGAGACCGAACAATAATATGTTCTAAT 240
```

```
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360

QY 361 TTTCACGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTCGAGTAGCC 420
Db 361 TTTCACGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTCGAGTAGCC 420

QY 421 ATAGCCCTTAGGATACAAAAGAAATTTTATCTTTCCGGAATTTGATTTTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTTAGGATACAAAAGAAATTTTATCTTTCCGGAATTTGATTTTTTATCAAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTTCTA 600

QY 601 GAAAAAATTTACAAAATAAATAAATACTATATGCTTATGCTCCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAATAAATACTATATGCTTATGCTCCTAATAGTCTTTTAGCAAAATTTT 660

QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTCAAAATTTTATCATACAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTCAAAATTTTATCATACAAGAAAAAATAACTACACT 720

QY 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780

QY 781 AAAAAAATAAATAAATAAAGAAATGTTTATACAAGTTGATATAAGACATCTATTAAAGATTA 840
Db 781 AAAAAAATAAATAAATAAAGAAATGTTTATACAAGTTGATATAAGACATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873

RESULT 11
US-10-303-118-13
; Sequence 13, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-118-13

Query Match      99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
   |||||
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60

QY 61 CTACCAAAATGATTTGATGATTAATCAATCAATTTTATTTTGAAGATAAATACTAT 120
   |||||
Db 61 CTACCAAAATGATTTGATGATTAATCAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTGTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATACTAC 180
   |||||
Db 121 CTGTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATACTAC 180

QY 181 ACTTTAAAAATTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240
   |||||
Db 181 ACTTTAAAAATTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTACGATTAATTTTCT 300
   |||||
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAAGAAATTTTAA 360
   |||||
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAAGAAATTTTAA 360

QY 361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGAGTAGCC 420
   |||||
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGAGTAGCC 420

QY 421 ATAGCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATCGGTC 480
   |||||
Db 421 ATAGCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATCGGTC 480

QY 481 TCTTATGCTTTTGTATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
   |||||
Db 481 TCTTATGCTTTTGTATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
   |||||
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTAATAGTCTTTTACGAAATTTT 660
   |||||
Db 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTAATAGTCTTTTACGAAATTTT 660

QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAATAAATACTACACT 720
   |||||
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAATAAATACTACACT 720

QY 721 AAAGATATCTCATACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
   |||||
Db 721 AAAGATATCTCATACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATATTAATTTT 780

QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTTATACAAAGTTGATAAAGACTTATTAAGATTA 840
   |||||
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTTATACAAAGTTGATAAAGACTTATTAAGATTA 840

QY 841 CCTAGTGATATAAAGCATTTTCAAGGAAAA 873
   |||||
Db 841 CCTAGTGATATAAAGCATTTTCAAGGAAAA 873
```

RESULT 12

US-10-303-128-13

; Sequence 13, Application US/10303128

```
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-128-13
```

```
Query Match      99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
   |||||
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60

QY 61 CTACCAAAATGATTTGATGATTAATCAATCAATTTTATTTTGAAGATAAATACTAT 120
   |||||
Db 61 CTACCAAAATGATTTGATGATTAATCAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTGTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATACTAC 180
   |||||
Db 121 CTGTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATACTAC 180

QY 181 ACTTTAAAAATTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240
   |||||
Db 181 ACTTTAAAAATTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTACGATTAATTTTCT 300
   |||||
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAAGAAATTTTAA 360
   |||||
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAAGAAATTTTAA 360

QY 361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGAGTAGCC 420
   |||||
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGAGTAGCC 420

QY 421 ATAGCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATCGGTC 480
   |||||
Db 421 ATAGCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATCGGTC 480

QY 481 TCTTATGCTTTTGTATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
   |||||
Db 481 TCTTATGCTTTTGTATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
   |||||
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
```

```
QY 601 GAAAAAAGCTTACAAAAATAAACTATATGCTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
DB 601 GAAAAAAGCTTACAAAAATAAACTATATGCTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAAATTTAAATTCAAATTTTATCATCAAGAAAAAATAAATTAACACT 720
DB 661 ATAGAACTAGCGCCAAAATTTAAATTCAAATTTTATCATCAAGAAAAAATAAATTAACACT 720
QY 721 AAAGATATACATACACTCTTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
DB 721 AAAGATATACATACACTCTTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAAATAAAAAATTTAAAGAAAATGTTTATTACAAAGTTGATAAAAGACTATTAAGATTA 840
DB 781 AAAAAAATAAAAAATTTAAAGAAAATGTTTATTACAAAGTTGATAAAAGACTATTAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTATTTTCAAAAGGAAAA 873
DB 841 CCTAGTGATATAAAGCATTATTTTCAAAAGGAAAA 873

RESULT 13
US-10-735-419-13
; Sequence 13, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-735-419-13
```

```
Query Match          99.3%; Score 869.8; DB 5; Length 873;
Best Local Similarity 99.8%; Pred No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAGAAAATTTGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAGAAAATTTGATTATTCAAGG 60
QY 61 CTACCAATATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAATATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTGTGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTTCAGCAATACTAC 180
DB 121 CTGTGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTTCAGCAATACTAC 180
QY 181 ACTTTTAAAAACATTTTAAATCCAAAATCAAGAATATGAGCCGAACACTAATTAATGTTCTAAT 240
DB 181 ACTTTTAAAAACATTTTAAATCCAAAATCAAGAATATGAGCCGAACACTAATTAATGTTCTAAT 240
```

```
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAGAAATTTAATGCTTATTTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAGAAATTTAATGCTTATTTTAAA 360
QY 361 TTTTCAGAAAATTTTATTTTCAATCAAGAAATTTACCTCAGGGGCTTATATGTCGAGTAGCC 420
DB 361 TTTTCAGAAAATTTTATTTTCAATCAAGAAATTTACCTCAGGGGCTTATATGTCGAGTAGCC 420
QY 421 ATAGCCCTTAGGATACAAAGAAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTTAGGATACAAAGAAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGGATACCAAAACAAAGAAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGGATACCAAAACAAAGAAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
DB 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAAATTTAAATTCAAATTTTATCATCAAGAAAAAATAAATTAACACT 720
DB 661 ATAGAACTAGCGCCAAAATTTAAATTCAAATTTTATCATCAAGAAAAAATAAATTAACACT 720
QY 721 AAAGATATACATCACTCTTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
DB 721 AAAGATATACATCACTCTTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAAATAAAAAATTTAAAGAAAATGTTTATTACAAAGTTGATAAAAGACTTATTAAAGATTA 840
DB 781 AAAAAAATAAAAAATTTAAAGAAAATGTTTATTACAAAGTTGATAAAAGACTTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTATTTTCAAAAGGAAAA 873
DB 841 CCTAGTGATATAAAGCATTATTTTCAAAAGGAAAA 873

RESULT 14
US-09-816-028A-11
; Sequence 11, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
```

US-09-816-028A-11

```
Query Match      98.9%; Score 866.6; DB 3; Length 873;
Best Local Similarity 99.5%; Pred. No. 3.7e-158;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTC AAGG 60
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTC AAGG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTTAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTTAA 360

QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
DB 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
DB 481 TCTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCCACTATATCGGACATGTAATAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATCGCTCCACTATATCGGACATGTAATAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATCAAAAATAAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840
DB 601 GAAAAAATCAAAAATAAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAGCATTTATTTC AAGGAAAA 873
DB 841 CCTAGTGATATAAGCATTTATTTC AAGGAAAA 873
```

RESULT 15

US-10-303-162-11

; Sequence 11, Application US/10303162

; Patent No. 6723545

; GENERAL INFORMATION:

; APPLICANT: Glibert, Michel

```
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-162-11
```

Query Match 98.9%; Score 866.6; DB 3; Length 873;

Best Local Similarity 99.5%; Pred. No. 3.7e-158;

Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTC AAGG 60
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTC AAGG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTTCTTTTGGAGCAATACTAC 180
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTTGGAGCAATACTAC 180

QY 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTTAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTTAA 360

QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
DB 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
DB 481 TCTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCCACTATATCGGACATGTAATAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATCGCTCCACTATATCGGACATGTAATAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATCAAAAATAAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 660
DB 601 GAAAAAATCAAAAATAAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 660
```

QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAGAAAATTAACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAGAAAATTAACACT 720  
QY 721 AAAGATATACATACCTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTAATTAATTT 780  
DB 721 AAAGATATACATACCTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTAATTAATTT 780  
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAAATCTATAAGATTA 840  
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAAATCTATAAGATTA 840  
QY 841 CCTAGTGATATAAAGCATTTATTCAAAGGAAAA 873  
DB 841 CCTAGTGATATAAAGCATTTATTCAAAGGAAAA 873

RESULT 16  
US-10-303-134-11  
; Sequence 11, Application US/10303134  
; Patent No. 6825019  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4

US-10-303-134-11

Query Match 98.9%; Score 866.6; DB 3; Length 873;  
Best Local Similarity 99.5%; Pred. No. 3.7e-158;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTCATTATTCACAGG 60  
DB 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTCATTATTCACAGG 60  
QY 61 CTACCAATGATTTTGTATGTTATAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGTATGTTATAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGACAGTTTATACACCCCTAAATTTCTTTTGGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGACAGTTTATACACCCCTGTTTCTTTCTTGGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATGAGACCGCAACTAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATGAGACCGCAACTAATATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACCATTTATTTCTCT 300  
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACCATTTATTTCTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
QY 361 TTTCAGCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420  
DB 361 TTTCAGCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTTCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTTCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600  
QY 601 GAAAAAATTTACAAAATAAAATCTATATGCTTATGCTTCTTATAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTTACAAAATAAAATCTATATGCTTATGCTTCTTACAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAGAAAATTAACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAGAAAATTAACACT 720  
QY 721 AAAGATATACATACCTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTAATTAATTT 780  
DB 721 AAAGATATACATACCTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTAATTAATTT 780  
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAGATTA 840  
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAGATTA 840  
QY 841 CCTAGTGATATAAAGCATTTATTCAAAGGAAAA 873  
DB 841 CCTAGTGATATAAAGCATTTATTCAAAGGAAAA 873

## RESULT 17

US-10-303-118-11  
; Sequence 11, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4

US-10-303-118-11

Query Match 98.9%; Score 866.6; DB 3; Length 873;  
Best Local Similarity 99.5%; Pred. No. 3.7e-158;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATAC 180  
DB 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCTTTTGAGCAATAC 180

QY 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240  
DB 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAATTAAGAAATTTAATGCTTTATTTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAATTAAGAAATTTAATGCTTTATTTTAAA 360

QY 361 TTTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGAATACCAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGAATACCAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTTCAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660  
DB 601 GAAAAAATTTTCAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660

QY 661 GAAAAAATTTTCAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 720  
DB 661 GAAAAAATTTTCAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 720

QY 721 AAAGATATACCTACCTTCTAGTGAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
DB 721 AAAGATATACCTACCTTCTAGTGAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780

QY 781 AAAAAATTTTAAAGAAATTTTATTTACAGTTGATAAAGATCTATTAAGATTA 840  
DB 781 AAAAAATTTTAAAGAAATTTTATTTACAGTTGATAAAGATCTATTAAGATTA 840

QY 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873  
DB 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 18  
US-10-303-128-11  
; Sequence 11, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.

APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-128-11

Query Match 98.9%; Score 866.6; DB 3; Length 873;  
Best Local Similarity 99.5%; Pred. No. 3.7e-158;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATAC 180  
DB 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCTTTTGAGCAATAC 180

QY 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240  
DB 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAATTAAGAAATTTAATGCTTTATTTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAATTAAGAAATTTAATGCTTTATTTTAAA 360

QY 361 TTTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGAATACCAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGAATACCAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTTCAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660  
DB 601 GAAAAAATTTTCAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660



QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAGAAAAAATAACTACACT 720  
|||||  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAGAAAAAATAACTACACT 720  
|||||  
QY 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAAATTT 780  
|||||  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAAATTT 780  
|||||  
QY 781 AAAAAATAAAAATTAAGAAAAATGTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 781 AAAAAATAAAAATTAAGAAAAATGTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
|||||  
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
|||||

RESULT 19  
US-10-735-419-11  
; Sequence 11, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-00011105  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-735-419-11

Query Match 98.9%; Score 866.6; DB 5; Length 873;  
Best Local Similarity 99.5%; Pred. No. 3.7e-158;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATTCATTATTCAAGG 60  
|||||  
Db 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATTCATTATTCAAGG 60  
|||||  
QY 61 CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
|||||  
Db 61 CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
|||||  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
|||||  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGTTTCTTTGAGCAATACTAC 180  
|||||  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATGAGCCGAACTAATATGTTCTTAAT 240  
|||||  
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATGAGCCGAACTAATATGTTCTTAAT 240  
|||||  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300  
|||||  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCT 300  
|||||

QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360  
|||||  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360  
|||||  
QY 361 TTTCACGAAATTTATTTCATCAAGAAATTAACCTCAGGGGCTATATGTCGCAAGTAGCC 420  
|||||  
Db 361 TTTCACGAAATTTATTTCATCAAGAAATTAACCTCAGGGGCTATATGTCGCAAGTAGCC 420  
|||||  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
|||||  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
|||||  
QY 481 TCTTATGCTTTTTCATACCAAGAAATCTTTTAAATCTTAGCCCTGATTTTAAAAAT 540  
|||||  
Db 481 TCTTATGCTTTTTCATACCAAGAAATCTTTTAAATCTTAGCCCTGATTTTAAAAAT 540  
|||||  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
|||||  
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
|||||  
QY 601 GAAAAAATTCAAAAATTAAGAAATATATGCTTATGCTTAAATGCTTTTACCAATTTT 660  
|||||  
Db 601 GAAAAAATTCAAAAATTAAGAAATATATGCTTATGCTTAAATGCTTTTACCAATTTT 660  
|||||  
QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATCTTATCATCAAGAAAAAATAACTACACT 720  
|||||  
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATCTTATCATCAAGAAAAAATAACTACACT 720  
|||||  
QY 721 AAAGATATACATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATATTAAATTT 780  
|||||  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATATTAAATTT 780  
|||||  
QY 781 AAAAAATAAAAATTAAGAAATATGTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 781 AAAAAATAAAAATTAAGAAATATGTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
|||||  
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
|||||

## RESULT 20

US-09-495-406-4  
; Sequence 4, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-0001105  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; CURRENT FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (CstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORP 7/a of LOS biosynthesis locus)  
US-09-495-406-4

Query Match 96.3%; Score 844; DB 3; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.2e-154;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
QY 1 ATGAAAAAGTTATTGCTGGAATGGCAAGTTTAAAGAAATGATTATTCAGG 60
Db |||||||
QY 1 ATGAAAAAGTTATTGCTGGAATGGCAAGTTTAAAGAAATGATTATTCAGG 60
Db |||||||
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTGAAGATAAATACTAT 120
Db |||||||
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTGAAGATAAATACTAT 120
Db |||||||
QY 121 CTTGGTAAAAAATGCAAGCAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
Db |||||||
QY 121 CTTGGTAAAAAATGCAAGCAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
Db |||||||
QY 181 ACTTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATCAATTTATGTTCTTAAT 240
Db |||||||
QY 181 ACTTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATCAATTTATGTTCTTAAT 240
Db |||||||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCCT 300
Db |||||||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCCT 300
Db |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAATTTTAAACAATTTTAAAGATTTTAAATTTTAAA 360
Db |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAATTTTAAACAATTTTAAAGATTTTAAATTTTAAA 360
Db |||||||
QY 361 TTTTCAGAAATTTATTCATCAAGAAATTTACCTCAGGAGTCTATATGTCGAGTAGCT 420
Db |||||||
QY 361 TTTTCAGAAATTTATTCATCAAGAAATTTACCTCAGGAGTCTATATGTCGAGTAGCT 420
Db |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCA 480
Db |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCA 480
Db |||||||
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTGGCTCCTGATTTTAAAT 540
Db |||||||
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTGGCTCCTGATTTTAAAT 540
Db |||||||
QY 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db |||||||
QY 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db |||||||
QY 601 GAAAAAATTTACAAAATAAACCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db |||||||
QY 601 GAAAAAATTTACAAAATAAACCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db |||||||
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720
Db |||||||
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720
Db |||||||
QY 721 AAAGATATCTCATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db |||||||
QY 721 AAAGATATCTCATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db |||||||
QY 781 AAAAAAATAAATAAAGAAATGTTTATTAAGATTTGATAAAGATCTATTAAGATTA 840
Db |||||||
QY 781 AAAAAAATAAATAAAGAAATGTTTATTAAGATTTGATAAAGATCTATTAAGATTA 840
Db |||||||
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAATAA 876
Db |||||||
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAATAA 876
Db |||||||
```

## RESULT 21

```
US-09-816-028A-4
; Sequence 4, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
```

```
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-09-816-028A-4
```

```
Query Match 96.3%; Score 844; DB 3; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.2e-154;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTGCTGGAATGGCAAGTTTAAAGAAATGATTATTCAGG 60
Db |||||||
QY 1 ATGAAAAAGTTATTGCTGGAATGGCAAGTTTAAAGAAATGATTATTCAGG 60
Db |||||||
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTGAAGATAAATACTAT 120
Db |||||||
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTGAAGATAAATACTAT 120
Db |||||||
QY 121 CTTGGTAAAAAATGCAAGCAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
Db |||||||
QY 121 CTTGGTAAAAAATGCAAGCAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
Db |||||||
QY 181 ACTTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATCAATTTATGTTCTTAAT 240
Db |||||||
QY 181 ACTTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATCAATTTATGTTCTTAAT 240
Db |||||||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCCT 300
Db |||||||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCCT 300
Db |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAATTTTAAACAATTTTAAAGATTTTAAATTTTAAA 360
Db |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAATTTTAAACAATTTTAAAGATTTTAAATTTTAAA 360
Db |||||||
QY 361 TTTTCAGAAATTTATTCATCAAGAAATTTACCTCAGGAGTCTATATGTCGAGTAGCT 420
Db |||||||
QY 361 TTTTCAGAAATTTATTCATCAAGAAATTTACCTCAGGAGTCTATATGTCGAGTAGCT 420
Db |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCA 480
Db |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCA 480
Db |||||||
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTGGCTCCTGATTTTAAAT 540
Db |||||||
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTGGCTCCTGATTTTAAAT 540
Db |||||||
QY 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db |||||||
QY 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db |||||||
QY 601 GAAAAAATTTACAAAATAAACCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db |||||||
QY 601 GAAAAAATTTACAAAATAAACCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db |||||||
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720
Db |||||||
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720
Db |||||||
```

```

Qy 721 AAAGATATATCTANACCTTCTAGTGAAGCTTATGGAAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAAGCTTATGGAAATTTTCAAAAAATATTAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAAAATGTTTATTACAAGTCATATAAAGATCTATTAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAAAATATTTTATTACAAGTCATATAAAGATCTATTAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 22
US-10-303-162-4
; Sequence 4, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (csfII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-4

Query Match 96.3%; Score 844; DB 3; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.2e-154;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATCAAAAAAGTTATTATTCGTGAATGACCAAGTTTAAAGAATTCGATTATTCACAGG 60
Db 1 ATCAAAAAAGTTATTATTCGTGAATGACCAAGTTTAAAGAATTCGATTATTCACAGG 60
Qy 61 CTACCAATGATTTTCATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTCATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATCTAC 180
Db 121 CTTGGTAAAAAATTCAAAGCAGTATTTTACAATCCTGGTCTTTTGTGAAACAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATAATATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATAATATATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTTGTAAAAACTTTTTTACGATTATTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTTGTAAAAACTTTTTTACGATTATTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTTAAAGAATTTAAATGCTTTATTA 360

```

Query Match 96.3%; Score 844; DB 3; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.2e-154;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTCCTGGAAATGACCAAGTTTAAAGAAATGATTTCAAGG 60  
DB 1 ATGAAAAAGTTATTTCCTGGAAATGACCAAGTTTAAAGAAATGATTTCAAGG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTTCGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTTCGAAGATAAATACTAT 120

QY 121 CTTCGTAATAAATGCAAGCAGTGTTCACACCCCTAAATTCCTCTTTGAGCAATACTAC 180  
DB 121 CTTCGTAATAAATGCAAGCAGTGTTCACACCCCTAAATTCCTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360

QY 361 TTTTCAGGAAATTTATCTCAATCAAGATATACCTCAGGAGTCTATATGTCAGTAGCT 420  
DB 361 TTTTCAGGAAATTTATCTCAATCAAGATATACCTCAGGAGTCTATATGTCAGTAGCT 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTAACAAATTAATGCTTATGTCCTTAATAGTCTTTTACGAAATTTT 660  
DB 601 GAAAAAATTAACAAATTAATGCTTATGTCCTTAATAGTCTTTTACGAAATTTT 660

QY 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAATAAATACTACACT 720  
DB 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAATAAATACTACACT 720

QY 721 AAAGATATACCTCATACCTCTAGTAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
DB 721 AAAGATATACCTCATACCTCTAGTAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780

QY 781 AAAAAATTAATAAATAAAGAAATGTTTATTAACAGTTGATAAAGATCTATTAAGATTA 840  
DB 781 AAAAAATTAATAAATAAAGAAATGTTTATTAACAGTTGATAAAGATCTATTAAGATTA 840

QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 876  
DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 876

## RESULT 24

US-10-303-118-4  
; Sequence 4, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-303-118-4

Query Match 96.3%; Score 844; DB 3; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.2e-154;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTCCTGGAAATGACCAAGTTTAAAGAAATGATTTCAAGG 60  
DB 1 ATGAAAAAGTTATTTCCTGGAAATGACCAAGTTTAAAGAAATGATTTCAAGG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTTCGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTTCGAAGATAAATACTAT 120

QY 121 CTTCGTAATAAATGCAAGCAGTGTTCACACCCCTAAATTCCTCTTTGAGCAATACTAC 180  
DB 121 CTTCGTAATAAATGCAAGCAGTGTTCACACCCCTAAATTCCTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360

QY 361 TTTTCAGGAAATTTATCTCAATCAAGATATACCTCAGGAGTCTATATGTCAGTAGCT 420  
DB 361 TTTTCAGGAAATTTATCTCAATCAAGATATACCTCAGGAGTCTATATGTCAGTAGCT 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTAACAAATTAATGCTTATGTCCTTAATAGTCTTTTACGAAATTTT 660  
DB 601 GAAAAAATTAACAAATTAATGCTTATGTCCTTAATAGTCTTTTACGAAATTTT 660

QY 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAATAAATACTACACT 720  
DB 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAATAAATACTACACT 720

QY 721 AAAGATATACCTCATACCTCTAGTAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
DB 721 AAAGATATACCTCATACCTCTAGTAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780

QY 781 AAAAAATTAATAAATAAAGAAATGTTTATTAACAGTTGATAAAGATCTATTAAGATTA 840  
DB 781 AAAAAATTAATAAATAAAGAAATGTTTATTAACAGTTGATAAAGATCTATTAAGATTA 840

QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 876  
DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 876



OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
; OTHER INFORMATION: biosynthesis locus)  
US-10-735-419-4

Query Match 96.3%; Score 844; DB 5; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.2e-154;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db |||||  
Qy 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db |||||  
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
Db |||||  
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
Db |||||  
Qy 121 CTTGGTAAAAATGCCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
Db |||||  
Qy 121 CTTGGTAAAAATGCCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
Db |||||  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAAATTTGTCTTAAT 240  
Db |||||  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAAATTTGTCTTAAT 240  
Db |||||  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCC 300  
Db |||||  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCC 300  
Db |||||  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
Db |||||  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
Db |||||  
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420  
Db |||||  
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420  
Db |||||  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGGTCA 480  
Db |||||  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAAATGATTTTATCAAAATGGGTCA 480  
Db |||||  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 540  
Db |||||  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 540  
Db |||||  
Qy 541 GATCGCTCGACCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Db |||||  
Qy 541 GATCGCTCGACCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Db |||||  
Qy 601 GAAAAAATTTTACAAAATAAATCTATATGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660  
Db |||||  
Qy 601 GAAAAAATTTTACAAAATAAATCTATATGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660  
Db |||||  
Qy 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATCAAGAAATAAATTAATCACTAC 720  
Db |||||  
Qy 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATCAAGAAATAAATTAATCACTAC 720  
Db |||||  
Qy 721 AAAGATATCTCATACCTCTAGTGAGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780  
Db |||||  
Qy 721 AAAGATATCTCATACCTCTAGTGAGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780  
Db |||||  
Qy 781 AAAAAAATAAATAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db |||||  
Qy 781 AAAAAAATAAATAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db |||||  
Qy 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATATA 876  
Db |||||  
Qy 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATATA 876  
Db |||||

RESULT 27

US-09-495-406-1

; Sequence 1, Application US/09495406

Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; CURRENT FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
; US-09-495-406-1

Query Match 95.6%; Score 837.6; DB 3; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.8e-152;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db |||||  
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
Db |||||  
Qy 6108 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 6167  
Db |||||  
Qy 121 CTTGGTAAAAATGCCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
Db |||||  
Qy 6168 CTTGGTAAAAATGCCAAGCAGTGTTTTACCAATCTATTTCTTTTGAACAACTAC 6227  
Db |||||  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAAATTTATGTTCTTAAT 240  
Db |||||  
Qy 6228 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAAATTTATGTTCTTAAT 6287  
Db |||||  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCC 300  
Db |||||  
Qy 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCC 6347  
Db |||||  
Qy 301 GATCCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
Db |||||  
Qy 6348 GATCCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 6407  
Db |||||  
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGAGTAGCC 420  
Db |||||  
Qy 6408 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGAGTAGCC 6467  
Db |||||  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTATTTTATCAAAATGGGTCA 480  
Db |||||  
Qy 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTATTTTATCAAAATGGGTCA 6527  
Db |||||  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 540  
Db |||||  
Qy 6528 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 6587  
Db |||||  
Qy 541 GATCGCTCGACCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600  
Db |||||  
Qy 6588 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 6647  
Db |||||  
Qy 601 GAAAAAATTTTACAAAATAAATCTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660  
Db |||||  
Qy 6648 GAAAAAATTTTACAAAATAAATCTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 6707  
Db |||||  
Qy 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATCAAGAAATAAATTAATCACTAC 720  
Db |||||

6708 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAAAATAACTACT 6767  
6721 AAAGATATACATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATATTATTT 780  
6768 AAAGATATACATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATATTATTT 6827  
781 AAAAAAATAAAAAATAAGAAAAATGTTTATTCAAGTTGATATAAGATCTATTAAAGATTA 840  
6828 AAAAAAATAAAAAATAAGAAAAATATTATTCAAGTTGATATAAGATCTATTAAAGATTA 6887  
841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876  
6888 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 6923

RESULT 28  
US-09-816-028A-1  
; Sequence 1, Application US/09816028A  
; Patent No. 6659705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
US-09-816-028A-1

Query Match 95.6%; Score 837.6; DB 3; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.8e-152;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

6048 ATGAAAAAGTTTATTATTCCTGCGAAATGACCAAGTTTAAAGAAATTCGATTATTTCAAGG 60  
6048 ATGAAAAAGTTTATTATTCCTGCGAAATGACCAAGTTTAAAGAAATTCGATTATTTCAAGA 6107  
61 CTACCAATGATTGATGATTATGATGTAATCAATCTTTTGAAGATAAATACTAT 120  
6108 CTACCAATGATTGATGATTATGATGTAATCAATCTTTTGAAGATAAATACTAT 6167  
121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTAATCTCTCTTGAGCAATACTAC 180  
6168 CTGTGTAATAAATGCAAGCAGTGTTTACAAATCCTTATCTTTTGAACAATACTAC 6227  
181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGCACTAATATGTTCTAAT 240  
6228 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGCACTAATATGTTCTAAT 6287  
241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGAAGAACTTTTACGATTATTTTCT 300  
6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGAAGAACTTTTACGATTATTTTCT 6347  
301 GATGCTCATTTGGGATATGATTTTTTAAACCACTTAAAGAAATTTAATGCTTATTTAAA 360  
6348 GATGCTCATTTGGGATATGATTTTTTCAACCACTTAAAGAAATTTAATGCTTATTTAAA 6407  
361 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGGCTATATGTCGAGTAGCC 420

6408 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTCGAGTAGCC 6467  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCA 480  
6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCA 6527  
481 TCTTATGCTTTTGCATACCAACCAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
6528 TCTTATGCTTTTGCATACCAACCAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 6587  
541 GATCGCTCGCACTATATCGGACATAGTFAAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
6588 GATAATTCACACTATATCGGACATAGTFAAAAAATACAGATATAAAGCTTTTGAATTTCTA 6647  
601 GAAAAAATTTACAAAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660  
6648 GAAAAAATTTACAAAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 6707  
661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATCAAGAAAAATAAATACTACT 720  
6708 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATCAAGAAAAATAAATACTACT 6767  
721 AAAGATATATCTATACCTTCTAGTGGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780  
6768 AAAGATATATCTATACCTTCTAGTGGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 6827  
781 AAAAAAATAAAAAATAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
6828 AAAAAAATAAAAAATAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887  
841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876  
6888 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 6923

RESULT 29  
US-10-303-162-1  
; Sequence 1, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
US-10-303-162-1

Query Match 95.6%; Score 837.6; DB 3; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.8e-152;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

6048 ATGAAAAAGTTTATTATTCCTGCGAAATGACCAAGTTTAAAGAAATTTAATGATTTCAAGG 60  
6048 ATGAAAAAGTTTATTATTCCTGCGAAATGACCAAGTTTAAAGAAATTTAATGATTTCAAGA 6107



```
QY 61 CTACCAATGATTTTGTATGTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 6108 CTACCAATGATTTTGTATGTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
QY 121 CTTGGTAAATAATGCAAGCAGTGTATTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180
Db 6168 CTTGGTAAATAATGCAAGCAGTGTATTTACAAATCCTATTCTTTTGTGAACAATACTAC 6227
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGCACTAATATGCTGCTTAAT 240
Db 6228 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGCACTAATATGCTGCTTAAT 6287
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTTCT 300
Db 6288 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTTCT 6347
QY 301 GATGCTCAITTTGGGATATGATTTTAAACCACTTTAAAGAAATTTAATGCTTTATTTAA 360
Db 6348 GATGCTCAITTTGGGATATGATTTTAAACCACTTTAAAGAAATTTAATGCTTTATTTAA 6407
QY 361 TTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db 6408 TTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 6467
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 6527
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 6528 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 6587
QY 541 GATGCTCGACATATACGACATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db 6588 GATAATTCACACTATATCGACATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTA 6647
QY 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTAGCAATTTT 660
Db 6648 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTAGCAATTTT 6707
QY 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAATAAATACTACACT 720
Db 6708 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAATAAATACTACACT 6767
QY 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780
Db 6768 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 6827
QY 781 AAAAAAATAAAATTAAGAAATGTTTATACAAGTTGATAAAAGATCTATTAAGATTA 840
Db 6828 AAAAAAATAAAATTAAGAAATGTTTATACAAGTTGATAAAAGATCTATTAAGATTA 6887
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 876
Db 6888 CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 6923
```

## RESULT 30

```
US-10-303-134-1
; Sequence 1, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
```

```
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-134-1
```

```
Query Match 95.6%; Score 837.6; DB 3; Length 11474;
Best Local Similarity 97.3%; Pred. No. 1.8e-152; Indels 0; Gaps 0;
Matches 852; Conservative 0; Mismatches 24;

QY 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
Db 6048 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAG 6107
QY 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 6108 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 6168 CTTGGTAAAAATGCAAGCAGTGTATTTTACAATCCTATTCTTTTGTGAACAATACTAC 6227
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAATATGTTGTTCTAAT 240
Db 6228 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAATATGTTGTTCTAAT 6287
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
Db 6288 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTTCT 6347
QY 301 GATGCTCATTTGGGATATGATTTTAAACCACTTTAAAGAAATTTAATGCTTTATTTAA 360
Db 6348 GATGCTCATTTGGGATATGATTTTAAACCACTTTAAAGAAATTTAATGCTTTATTTAA 6407
QY 361 TTTCACCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db 6408 TTTCACCAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 6467
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 6527
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATTTT 540
Db 6528 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATTTT 6587
QY 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db 6588 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 6647
QY 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTAGCAATTTT 660
Db 6648 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTAGCAATTTT 6707
QY 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAATAAATACTACACT 720
Db 6708 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAATAAATACTACACT 6767
QY 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780
Db 6768 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 6827
QY 781 AAAAAAATAAAATTAAGAAATGTTTATACAAGTTGATAAAAGATCTATTAAGATTA 840
Db 6828 AAAAAAATAAAATTAAGAAATGTTTATACAAGTTGATAAAAGATCTATTAAGATTA 6887
```







```
US-09-495-406-2
; Sequence 2, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of LOS biosynthesis locus)
US-09-495-406-2

Query Match          95.4%; Score 836; DB 3; Length 876;
Best Local Similarity 97.1%; Pred. No. 2.9e-152;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTCATTATTCGAAG 60
DB 1 ATGAAAAAGTTATTATTCGGAATGACCAAGTTTAAAGAAATTCATTATTCGAAG 60
QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTGGTAAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
DB 121 CTGGTAAAAAATGCAAGCAGTGTTTACAAATCCTATTCTTTTTTGAACAATACTAC 180
QY 181 ACTTTAAAAACATTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
DB 181 ACTTTAAAAACATTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
QY 241 TACCAACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACTTTTTACCAGTATTATTTTCCT 300
DB 241 TACCAACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACTTTTTACCAGTATTATTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 361 TTTCAGAAATTTATTTCAATCAAGAAATTAATCTAGGGGTCTATATGTCAGTAGCC 420
DB 361 TTTCAGAAATTTATTTCAATCAAGAAATTAATCTAGGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTTAAAAATAGCCCTCATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACTAAAACAAAAATCTTTTAAAAATGGCTCCTTAATTTAAAAAT 540
QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600
DB 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAACTTACAAAATAAAACTATATTCGTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
DB 601 GAAAAAACTTACAAAATAAAACTATATTCGTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATCAAGAAATAAATACTACTAC 720
DB 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATCAAGAAATAAATACTACTAC 720
QY 721 AAAGATATACTCATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
DB 721 AAAGATATACTCATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAAATAA 876
DB 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAAATAA 876

RESULT 35
US-09-816-028A-2
; Sequence 2, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-09-816-028A-2

Query Match          95.4%; Score 836; DB 3; Length 876;
Best Local Similarity 97.1%; Pred. No. 2.9e-152;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTCATTATTCGAAG 60
DB 1 ATGAAAAAGTTATTATTCGGAATGACCAAGTTTAAAGAAATTCATTATTCGAAG 60
QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTGGTAAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
DB 121 CTGGTAAAAAATGCAAGCAGTGTTTACAAATCCTATTCTTTTTTGAACAATACTAC 180
QY 181 ACTTTAAAAACATTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
DB 181 ACTTTAAAAACATTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
QY 241 TACCAACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACTTTTTACCAGTATTATTTTCCT 300
DB 241 TACCAACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACTTTTTACCAGTATTATTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 361 TTTCAGAAATTTATTTCAATCAAGAAATTAATCTAGGGGTCTATATGTCAGTAGCC 420
DB 361 TTTCAGAAATTTATTTCAATCAAGAAATTAATCTAGGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTTAAAAATAGCCCTCATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACTAAAACAAAAATCTTTTAAAAATGGCTCCTTAATTTAAAAAT 540
QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600
DB 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAACTTACAAAATAAAACTATATTCGTTATGTCCTTAATAGTCTTTTAGCAAAATTTT 660
```

```
Db 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAAACCTTTTACGATTATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAA 360
Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATATACCAAAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATATACCAAAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATAATTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTCAAAAATAAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTCAAAAATAAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTTATCAATCAAGAAATTTTATCAAGAAATTAACACACT 720
Db 661 ATAGAACTAGCCCAAAATTTAAATTTTATCAATCAAGAAATTTTATCAAGAAATTAACACACT 720
Qy 721 AAAGATATATCTCATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATTAATTTT 780
Db 721 AAAGATATATCTCATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATTAATTTT 780
Qy 781 AAAAAAATAAAATAAAGAAATGTTTATACAGTTGATATAAGAGATCTATTAAGATTA 840
Db 781 AAAAAAATAAAATAAAGAAATGTTTATACAGTTGATATAAGAGATCTATTAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
```

## RESULT 36

```
US-10-303-162-2
; Sequence 2, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
```

```
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-2
```

```
Query Match 95.4%; Score 836; DB 3; Length 876;
Best Local Similarity 97.1%; Pred. No. 2.9e-152;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60
Db 1 ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGA 60
Qy 61 CTACCAATCATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATCATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTCTTTTGAGCAATACATAC 180
Db 121 CTTGGTAAAAATGCAAGGCAGTATTTTACAATCTATTTCTTTTGTGAACAATACATAC 180
Qy 181 ACTTTAAACATTTAATCCAAAATCAAGATATAGACCCGAACCTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGATATAGACCCGAACCTAATTTATGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTAAAAAATTTTACGATTTATTTCTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTAAAAAATTTTACGATTTATTTCTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTTAAATGCTTTATTTAAA 360
Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATATCAAAACAAGAAATCTTTTAAAAATCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATATCAAAACAAGAAATCTTTTAAAAATCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600
Db 541 GATAATTCACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTCAAAAATAAATACTATATGCTTATGCTTCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTCAAAAATAAATACTATATGCTTATGCTTCTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTTATCAATCAAGAAATTTTATCAAGAAATTAACACACT 720
Db 661 ATAGAACTAGCCCAAAATTTAAATTTTATCAATCAAGAAATTTTATCAAGAAATTAACACACT 720
Qy 721 AAAGATATATCTCATACCTCTAGTAGGCTTTATGGAATTTTCAAAAATATTAATTTT 780
Db 721 AAAGATATATCTCATACCTCTAGTAGGCTTTATGGAATTTTCAAAAATATTAATTTT 780
Qy 781 AAAAAAATAAAATAAAGAAATGTTTATTAAGAGATTTTACAGTTGATATAAGAGATCTATTAAGATTA 840
Db 781 AAAAAAATAAAATAAAGAAATGTTTATTAAGAGATTTTACAGTTGATATAAGAGATCTATTAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
```

## RESULT 37

```
US-10-303-134-2
; Sequence 2, Application US/10303134
```

Patent No. 6825019  
GENERAL INFORMATION:  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/303,134  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
OTHER INFORMATION: biosynthesis locus)  
US-10-303-134-2

Query Match 95.4%; Score 836; DB 3; Length 876;  
Best Local Similarity 97.1%; Pred. No. 2.9e-152;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAG 60  
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTCTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGTGAACAATACTAC 180  
QY 181 ACTTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAATTTATGTTCTTAAT 240  
DB 181 ACTTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAATTTATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCTCT 300  
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACTCAGGGGCTATATGTTGTCAGTAGCC 420  
DB 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACTCAGGGGCTATATGTTGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGTATACCAACAAAGAAATTTCTTTTAAATCTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTGTATACCAACAAAGAAATTTCTTTTAAATCTAGCCCTGATTTTAAAT 540  
QY 541 GATCGCTCGCATATATCGCATAGTAAATACAGATATAAAGCTTTAGAAATTTCTTA 600

DB 541 GATAATTTACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAATTAATACTATATTGCTTAATGCTTAATAGTCTTTTAGCAAAATTTT 660  
DB 601 GAAAAAATTTACAAATTAATACTATATTGCTTAATGCTTAATAGTCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCAAAATTTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATCTACACT 720  
DB 661 ATAGAACTAGCGCAAAATTTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATCTACACT 720  
QY 721 AAAGATATATCTCATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATTTTAAATTTT 780  
DB 721 AAAGATATATCTCATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATTTTAAATTTT 780  
QY 781 AAAAAATTAATAAATTAAGAAATTTTATACAGTTGATAAAGATCTTATTAAGATTA 840  
DB 781 AAAAAATTAATAAATTAAGAAATTTTATTAATTTTATACAGTTGATAAAGATCTTATTAAGATTA 840  
QY 841 CCTAGTGATATAAAGCATTTTCAAAAGGAAATAA 876  
DB 841 CCTAGTGATATAAAGCATTTTCAAAAGGAAATAA 876

RESULT 38  
US-10-303-118-2  
Sequence 2, Application US/10303118  
Patent No. 690867  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/303,118  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
OTHER INFORMATION: biosynthesis locus)  
US-10-303-118-2

Query Match 95.4%; Score 836; DB 3; Length 876;  
Best Local Similarity 97.1%; Pred. No. 2.9e-152;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60  
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTCTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGTGAACAATACTAC 180

181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTTAAT 240  
181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTTAAT 240  
241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300  
241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300  
301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTTTAAAA 360  
301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTTTAAAA 360  
361 TTTTACGAAATTTAATTTCAATCAAGAAATTAACCTCAGGGCTCTATATGTCGAGTAGCC 420  
361 TTTTACGAAATTTAATTTCAATCAAGAAATTAACCTCAGGGCTTTATATGTCGAGTAGCC 420  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
481 TCTTATGCTTTTGCATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
481 TCTTATGCTTTTGCATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
541 GATGCTCGCACTATATCGACATAGTAAACAAATACAGATATATAAGGCTTTAGAAATTTCTA 600  
541 GATGCTCGCACTATATCGACATAGTAAACAAATACAGATATATAAGGCTTTAGAAATTTCTA 600  
601 GAAAAAATTTACAAATTAACCTATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660  
601 GAAAAAATTTACAAATTAACCTATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660  
661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATATCAAGAAATTAACCTAGTCT 720  
661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATATCAAGAAATTAACCTAGTCT 720  
721 AAAGATATATCTATCTTCTAGTGGCTTATGGAATTTTCAAAATATTAATTTT 780  
721 AAAGATATATCTATCTTCTAGTGGCTTATGGAATTTTCAAAATATTAATTTT 780  
781 AAAAAATTAATAATTAAGAAATTTTATTAACAAAGTTGATAAAGATCTATTAAAGATTA 840  
781 AAAAAATTAATAATTAAGAAATTTTATTAACAAAGTTGATAAAGATCTATTAAAGATTA 840  
841 CCTAGTATATAAGCATTTTCAAGGAAATAA 876  
841 CCTAGTATATAAGCATTTTCAAGGAAATAA 876

RESULT 39

US-10-303-128-2  
; Sequence 2, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA

ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))  
OTHER INFORMATION: biosynthesis locus  
US-10-303-128-2

Query Match 95.4%; Score 836; DB 3; Length 876;  
Best Local Similarity 97.1%; Pred. No. 2.9e-152;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTTTCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTTTCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 60  
61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
121 CTTGGTAAAAATGCAAGCAGTGTTTTACAAATCCTATTTCTTTTGAACAATACTAC 180  
181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACTAATATGTTGTTCTAAT 240  
181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACTAATATGTTGTTCTAAT 240  
241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTTCTCT 300  
241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTTCTCT 300  
301 GATGCTCATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAAATGCTTTTAAAA 360  
301 GATGCTCATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTTAAATGCTTTTAAAA 360  
361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGCTCTATATGTCGAGTAGCC 420  
361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGCTTTATATGTCGAGTAGCC 420  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
541 GATGCTCGCACTATATCGGACATAGTAAACAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
541 GATGCTCGCACTATATCGGACATAGTAAACAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
601 GAAAAAATTTACAAATTAACCTATATGCTTATGCTTCTTAATAGTCTTTTACCAATTTT 660  
601 GAAAAAATTTACAAATTAACCTATATGCTTATGCTTCTTAATAGTCTTTTACCAATTTT 660  
661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATATCAAGAAATTAACCTAGTCT 720  
661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATATCAAGAAATTAACCTAGTCT 720  
721 AAAGATATATCTATCTTCTAGTGGCTTTATGGAATTTTCAAAATATTAATTTT 780  
721 AAAGATATATCTATCTTCTAGTGGCTTTATGGAATTTTCAAAATATTAATTTT 780  
781 AAAAAATTAATAATTAAGAAATTTTATTAACAAAGTTGATAAAGATCTATTAAAGATTA 840  
781 AAAAAATTAATAATTAAGAAATTTTATTAACAAAGTTGATAAAGATCTATTAAAGATTA 840  
841 CCTAGTATATAAGCATTTTCAAGGAAATAA 876  
841 CCTAGTATATAAGCATTTTCAAGGAAATAA 876

RESULT 40  
 US-10-735-419-2  
 ; Sequence 2, Application US/10735419  
 ; Patent No. 7026147  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/735,419  
 ; CURRENT FILING DATE: 2003-12-11  
 ; PRIOR APPLICATION NUMBER: US/09/816,028A  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
 ; OTHER INFORMATION: biosynthesis locus)  
 US-10-735-419-2

Query Match	95.4%	Score 836;	DB 5;	Length 876;
Best Local Similarity	97.1%;	Pred. No. 2.9e-152;		
Matches 851;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
Qy 1	ATGAAAAAGATTATTATTCTCGAATGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60			
Db 1	ATGAAAAAGATTATTATTCTCGAATGACCAAGTTTAAAAAGAAATGATTATTCAAGA 60			
Qy 61	CTACCAAAATGATTTTCATGCTATTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120			
Db 61	CTACCAAAATGATTTTCATGCTATTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120			
Qy 121	CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180			
Db 121	CTTGGTAAAAATGCAAGCAGTATTTTTACAATCCCTATTCTTTTTTTTGAACAATACTAC 180			
Qy 181	ACTTTAAAAACATTTAATCCAAATCAAGAATATGACCGCACTAATATATGTGTTCTAAT 240			
Db 181	ACTTTAAAAACATTTAATCCAAATCAAGAATATGACCGCACTAATATATGTGTTCTAAT 240			
Qy 241	TACAACCAAGCTCATCTGAAAAATGAAAAATTTTGTAAAAACCTTTTATTCAGATTATTTTCT 300			
Db 241	TACAACCAAGCTCATCTGAAAAATGAAAAATTTTGTAAAAACCTTTTATTCAGATTATTTTCT 300			
Qy 301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTTAAAGAAATTAATGCTTATTTAAA 360			
Db 301	GATGCTCATTTGGGATATGATTTTTTTTCAACAACCTTTAAAGATTTTAAATGCTTATTTAAA 360			
Qy 361	TTTCAGGAAATTTATTTTCAATCAAGAAATTAATCTCAGGGGCTCTATATGTGTCAGTAGCC 420			
Db 361	TTTCAGGAAATTTATTTTCAATCAAGAAATTAATCTCAGGGGTTTATATGTGTCAGTAGCC 420			
Qy 421	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAATGGTCA 480			
Db 421	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAATGGTCA 480			
Qy 481	TCTTATGCTTTTGATACCAACCAAGAAAAATCTTTTTAAAACTAGCCCTGATTTTAAAAAT 540			

```
Db 181 ACTTTAAACATTTAAATCCAAATCAAGAAATAGAGACCAGAACTAAATCATGTGCTCTAAT 240
Qy 241 TACAACNAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTTCCT 300
Db 241 TTTAAACNAGCTCATCTAGAAATCAAAATTTTGTAAACCTTTTACGATTAATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATGATTTTATCAAAATGGGATCA 480
Qy 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAACATACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATTAATTCACACTATATCGGACATAGTAAACATACAGATATATAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAACTATATGTCTTATGTCTCTAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACGAAATAAAGCTATATGTTTATGTCTTAAACAGTCTTTTACGAAATTTT 660
Qy 720 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAGAAATAAATACTACACT 720
Db 720 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAGAAATAAATACTATACT 720
Qy 780 AAAGATATATCTATACCTCTTAGTAGGCTTTATGTTGAAATTTTCAAAATAATTTAAATTT 780
Db 780 AAAGATATATCTATACCTCTTAGTAGGCTTTATGTTGAAATTTTCAAAATAATTTAAATTT 780
Qy 840 AAAAAATAAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTA 840
Db 840 AAAAAATAAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTA 840
Qy 841 CCTAGTGATTAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATTAAGCAATTTTCAAGGAAATAA 876
```

## RESULT 42

```
US-09-816-028A-6
; Sequence 6, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)...(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-09-816-028A-6
```

```
Query Match 93.6%; Score 820; DB 3; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.4e-149;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTTATTTATGCTGGAAATCGAACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTTATGCTGGAAATCGAACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60
Qy 61 CTACCAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAAATGCAAAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAAATGCAAAAGCAGTATTTTACAAATCTAGTCTTTTGTGAAACAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAAATCATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTTATTTCTCT 300
Db 241 TTTAACCAAGCTCATCTAGAAAAATCAAAATTTTGTAAAAAATCTTTTACGATTTATTTCTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTCAATGCTTTATTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAATGGATCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAATGGATCA 480
Qy 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAATTAATTTTCAAAATTTTATCATACAGAAATAAATACTACACT 720
Db 601 GAAAAAATTTACGAAATAAAGCTATATGTTTATGTCTTAAACAGTCTTTTACGAAATTTT 720
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAATAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAATAAATACTATACT 720
Qy 721 AAAGATATATCTATACCTCTTAGTAGGCTTTATGTTGAAATTTTCAAAATAATTTAAATTT 780
Db 721 AAAGATATATCTATACCTCTTAGTAGGCTTTATGTTGAAATTTTCAAAATAATTTAAATTT 780
Qy 781 AAAAAATAAAAAATTAAGAAATGTTTATTTTACAAGTTGATAAAGATCTTATTAAGATTA 840
Db 781 AAAAAATAAAAAATTAAGAAATGTTTATTTTACAAGTTGATAAAGATCTTATTAAGATTA 840
Qy 841 CCTAGTGATTAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATTAAGCAATTTTCAAGGAAATAA 876
```

## RESULT 43

```
US-10-303-162-6
; Sequence 6, Application US/10303162
```



```
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype 0:41
; US-10-303-162-6

Query Match          93.6%; Score 820; DB 3; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.4e-149;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGG 60
DB      1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGA 60

QY      61 CTACCAATGATTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120
DB      61 CTACCAATGATTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120

QY      121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
DB      121 CTTGGTAAAAATGCAAGCAGTGATTTTACATCCTAGTCTTTTTTGAACAATACTAC 180

QY      181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATTAATGTTCTAAT 240
DB      181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATCAATGTTCTAAT 240

QY      241 TCAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCT 300
DB      241 TTTAACCAAGCTCATCTAGAAATCAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

QY      301 GATGCTCATTGGGATATGATTTTAAACACTTTAAAGAAATTTAATGCTTTATTTAA 360
DB      301 GATGCTCATTGGGATATGATTTTCAACCAACTTTAAAGAAATCAATGCTTTATTTAA 360

QY      361 TTTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
DB      361 TTTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

QY      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGGTCA 480
DB      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGATCA 480

QY      481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
DB      481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAATTTGCTCTCTAATTTTAAAT 540

QY      541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB      541 GATAAATTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

105 601 GAAAAAACTTACAAAAATAAACTATATGCTTCTAATAGTCTTTTAGCAAAATTTT 660
106 601 GAAAAAACTTACGAAATAAAGCTATATGTTTATGCTTAACAGTCTTTTAGCAAAATTTT 660
107 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720
108 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACTAT 720
109 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780
110 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780
111 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
112 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
113 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876
114 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876

RESULT 44
US-10-303-134-6
; Sequence 6, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype 0:41
; US-10-303-134-6

Query Match          93.6%; Score 820; DB 3; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.4e-149;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGG 60
DB      1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGA 60

QY      61 CTACCAATGATTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120
DB      61 CTACCAATGATTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120

QY      121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
DB      121 CTTGGTAAAAATGCAAGCAGTGATTTTACATCCTAGTCTTTTTTGAACAATACTAC 180

QY      181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATTAATGTTCTAAT 240
DB      181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACATAATCAATGTTCTAAT 240

QY      241 TCAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCT 300
DB      241 TTTAACCAAGCTCATCTAGAAATCAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

QY      301 GATGCTCATTGGGATATGATTTTAAACACTTTAAAGAAATTTAATGCTTTATTTAA 360
DB      301 GATGCTCATTGGGATATGATTTTCAACCAACTTTAAAGAAATCAATGCTTTATTTAA 360

QY      361 TTTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
DB      361 TTTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

QY      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGGTCA 480
DB      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGATCA 480

QY      481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
DB      481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAATTTGCTCTCTAATTTTAAAT 540

QY      541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB      541 GATAAATTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
```

```
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
Db 241 TTTTAAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTTAAAGAAATTTAATGCTTAATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTTTCAACAACCTTTAAAGAAATTTAATGCTTAATTTTAAA 360
QY 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGCTGCGAGTAGCC 420
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGCTGCGAGTAGCC 420
QY 421 ATAGCCTTAGGATACAAAGAAATTTTATCTTTTCGGGAATTCATTTTTATCAAAATGGGTCA 480
Db 421 ATAGCCTTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACAAAGAAATTTTATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTATCTTTTAAATTTGGCTCTAAATTTTAAAAAT 540
QY 541 GATGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAAATTTCTA 600
Db 541 GATAATTTCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTTACAAATTAATAATTTTATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTTACAAATTAATAATTTTATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATCAAGAAATAAATACTACACT 720
Db 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATCAAGAAATAAATACTACACT 720
QY 721 AAAGATATATCTATACCTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATATCTATACCTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATTTAATTTT 780
QY 781 AAAAAATTAATAATTAAGAAATGTTTATTACAAGTTGATAAAGACTCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATGTTTATTACAAGTTGATAAAGACTCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876
```

## RESULT 45

```
US-10-303-118-6
; Sequence 6, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
```

```
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-303-118-6
```

```
Query Match 93.6%; Score 820; DB 3; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.4e-149;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAAGTTATTATTGCTGGAAATGCAACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
Db 1 ATGAAAAAAGTTATTATTGCTGGAAATGCAACCAAGTTTAAAAAGAAATTTGATTATTCAAGA 60
QY 61 CTACCAATGATTTTGTATGTTTATAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTTGTATGTTTATAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
QY 121 CTTGGTAAAAAATGCAAGCAAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAAAATCTAC 180
Db 121 CTTGGTAAAAAATGCAAGCAAGTGTTTTACAAATCCTAGTCTTTTTTTTGAACAATCTAC 180
QY 181 ACTTTAAACATTTTAAATCCAAAAATCAAGAATATGAGCCGAACATAATATATGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAAAATCAAGAATATGAGCCGAACATAATCAATGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTATCGATTTATTTCT 300
Db 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTATCGATTTATTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTTAAAGAAATTTTAAATGCTTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTTCAAACAACCTTTAAAGAAATTTCAATGCTTTTAAA 360
QY 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAATGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTTTATCAAAATGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAATTTTAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAATTTTAAATTTTAAATTTGGCTCTAAATTTTAAAAAT 540
QY 541 GATGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATAATTTCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTTACAAATAAATACTATATGCTTATGCTCTAATAGTCTTTTACCAATTTT 660
Db 601 GAAAAAATTTTACGAAATAAAGCTATATTTTATGCTCTACAGTCTTTTACCAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATCAAGAAATAAATACTACACT 720
Db 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATCAAGAAATAAATACTACACT 720
QY 721 AAAGATATATCTATACCTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATATCTATACCTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATTTAATTTT 780
QY 781 AAAAAATTAATAATTAAGAAATGTTTATTACAAGTTGATAAAGACTCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATGTTTATTACAAGTTGATAAAGACTCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876
```

## RESULT 46

```
US-10-303-128-6
; Sequence 6, Application US/10303128
; Patent No. 6911337
```

```
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
; US-10-303-128-6

Query Match          93.6%; Score 820; DB 3; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.4e-149;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCGCTGGAAGTGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60
DB 1 ATGAAAAAGTTATTATTCGCTGGAAGTGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60
QY 61 CTACCAAAATGATTTGATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120
DB 61 CTACCAAAATGATTTGATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
DB 241 TACACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCCT 300
QY 301 GATGCTCATTGTGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360
DB 301 GATGCTCATTGTGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360
QY 361 TTTACAGAAATTTAATTCGAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
DB 361 TTTACAGAAATTTAATTCGAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTACAAAAATAAACTATATGTGTTAAGTTCCTAAATAGTCTTTTAGCAAAATTTT 660
DB 601 GAAAAAATTTACGAAATAAAGCTATATTTGTTTGTCTTAACAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAATTTTATCATACAGAAAAAATAACTACTACT 720
DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAATTTTATCATACAGAAAAAATAACTACTACT 720
QY 721 AAAGATATATCTCATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780
DB 721 AAAGATATATCTCATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780
QY 781 AAAAAAATAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTTATTAAAGATTA 840
DB 781 AAAAAAATAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 876
DB 841 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 876

RESULT 47
US-10-735-419-6
; Sequence 6, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
; US-10-735-419-6

Query Match          93.6%; Score 820; DB 5; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.4e-149;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCGCTGGAAGTGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60
DB 1 ATGAAAAAGTTATTATTCGCTGGAAGTGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60
QY 61 CTACCAAAATGATTTGATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120
DB 61 CTACCAAAATGATTTGATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAATGAAAAATTTGTAAAAACTTTTACGATTATTTTCCT 300
DB 241 TTAACCAAGCTCATCTAGAAAATGAAAAATTTGTAAAAACTTTTACGATTATTTTCCT 300
QY 301 GATGCTCATTGTGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360
DB 301 GATGCTCATTGTGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360
QY 361 TTTACAGAAATTTAATTCGAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
DB 361 TTTACAGAAATTTAATTCGAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTACAAAAATAAACTATATGTGTTAAGTTCCTAAATAGTCTTTTAGCAAAATTTT 660
```

```
Db 241 TTTAAACCAAGCTCATCTAGAAAATCAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAAATGCTTTATTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACCTTAAAGAAATTTAAATGCTTTATTTAA 360
Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAATTTGATTTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAATTTGATTTTTATCAAAATGGATCA 480
Qy 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTTTAAAT 540
Db 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAATTTGGCTCCTAATTTTTTAAAT 540
Qy 541 GATCGCTCGCATATATCGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600
Db 541 GATAATTCACACTATATCGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600
Qy 601 GAAAAAATTTCAAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTCAAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAATAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAATAAATACTACTAT 720
Qy 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATATATTTT 780
Db 721 AAAGATATATCTATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATATATTTT 780
Qy 781 AAAAAAATAAAATTTAAAGAAATTTTATTTACAAGTTGATAAAGATCTATTAAGATTA 840
Db 781 AAAAAAATAAAATTTAAAGAAATTTTATTTACAAGTTGATAAAGATCTATTAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
```

## RESULT 48

```
US-09-272-960-1
; Sequence 1, Application US/09272960
; Patent No. 6689604
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Lipopolysaccharide alpha-2,3 sialyltransferase of
; TITLE OF INVENTION: Campylobacter jejuni and Its Uses
; FILE REFERENCE: 014137-013210US
; CURRENT APPLICATION NUMBER: US/09/272,960
; CURRENT FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: US 60/078,891
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: US 09/272,960
; EARLIER FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1293)
; OTHER INFORMATION: Campylobacter jejuni OH4384 cst-1 gene
; OTHER INFORMATION: alpha2,3-sialyltransferase
US-09-272-960-1
```

```
Query Match 35.6%; Score 311.6; DB 3; Length 1293;
Best Local Similarity 61.1%; Pred. No. 2.7e-51;
Matches 522; Conservative 0; Mismatches 329; Indels 3; Gaps 1;

Qy 1 ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60
Db 46 ATGCAAAATATAATCATAGCAGGAAATGGACCTAGCCTAAAAAATATTTAAATTTAAAAA 105
Qy 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 106 CTGCCTAGAGAAATATGATTTTATAGGTGTAAACAGTTTATTTTGAAGATAAGTATTAT 165
Qy 121 CTTGGTAAAAAATGCAAAAGCAGTCTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
Db 166 TTAGAAAAAAGATTAAGACGATTTTAAATCTCGGTGCTTTTACACAGTATCAC 225
Qy 181 ACTTTAAAAACATTTAAATCCAAAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCT 240
Db 226 ACTGCAAAACCAACTTATCTACTAAAAAATGAGTATGAAATAAAAAATATTTTGTCTCTACA 285
Qy 241 TACAAACAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300
Db 286 TTTAAATTTACCTTTTATTTGAAGCAATGATTTTATACATCAATTTTATTAATTTTTCGCC 345
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAAAACAACCTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 346 GATGCAAAACCTTGCTGTATGAAGTTTATTTGAAAACTTTAAAGAAATTTTATGCTTTATATA 405
Qy 361 TTTTCAAGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db 406 TACAATGAATTTATTTCAATTAAGAAATTTACTTCGGSGCTATATATGTCGCAATTTGCT 465
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 466 ATTGCATTAGGATATAAAACCATCTATTTATGTGGCATTTGATTTTATGAGGAGATGTT 525
Qy 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 526 ATTTATCCTTTTGAAGCTATGAGTACAAATATAAAACAATCTTTCTGGGAATAAAA --- 582
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db 583 GATTTCAAACTTCAAAATTTGTCATTTAAGAAATACGATATAGAAGCATTAATAATTTGTTA 642
Qy 601 GAAAAAATTTCAAAAAATAAACTATATTTGCTTATGTCCTAATAGTCTTTTACCAATTTT 660
Db 643 AAATCAATATACAAAGTTAATATCTACGCAATTTGTGTGATGATTTCTATTTTGGCAATCAT 702
Qy 661 ATAGAACTAGGCCAAATTTTAAATTTCAATTTTATCATCAAGAAAAATAACTACACT 720
Db 703 TTTCTTTTATCAATTTAATAATTAATTAACAAATTTCACTTTTAGAAAAATAGCATTAATTTCT 762
Qy 721 AAAGATATACTCATACCTCTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTTAATTTT 780
Db 763 ATAAATGATATTTTATTTGACGATATACTCTCGCGGTAGTTTTATATAAAATCAACTT 822
Qy 781 AAAAAAATAAAATTTAAAGAAATTTTATTTACAAGTTTGAATAAAGATCTATTAAGATTA 840
Db 823 AAAGCTGATAATAAAATTTATGCTTTAATTTTATAATTTCTTCTTAAAGATAATTTA 882
Qy 841 CCTAGTGATATAA 854
Db 883 ATTAATTTTAAA 896
```

## RESULT 49

```
US-10-058-636-1
; Sequence 1, Application US/10058636
; Patent No. 6709834
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
```

```

; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Lipopolysaccharide alpha-2,3 Sialyltransferase of
; TITLE OF INVENTION: Campylobacter jejuni and Its Uses
; FILE REFERENCE: 014137-013210US
; CURRENT APPLICATION NUMBER: US/10/058,636
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/09/272,960
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: US 60/078,891
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/272,960
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1293)
; OTHER INFORMATION: Campylobacter jejuni OH4384 cst-I gene
; OTHER INFORMATION: alpha2,3-sialyltransferase
US-10-058-636-1

```

Query Match	35.6%	Score 311.6	DB 3	Length 1293	
Best Local Similarity	61.1%	Pred. No. 2.7e-51			
Matches 522	Conservative	0	Mismatches 329	Indels 3	Gaps 1
QY	1	ATGAAAAAGCTTATTATTCGTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCGAAGG	60		
DB	46	ATGCAAAATATAATCATAGCAGGAATGGACCTTAGCCTAAATAATTAATTATAAAGA	105		
QY	61	CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTTATTTTGAAGATAAATACTAT	120		
DB	106	CTGCCTAGAGAAATATGATGTTTTTAGGTGAACCCAGTTTTATTTTGAAGATAAGTATTAT	165		
QY	121	CTTGGTAAAAAATGCAAAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC	180		
DB	166	TTAGGAAAAAAGATATAAGCAGTATTTTTTAAATCCTGGTGCTTTTTTACAAACAGTATCAC	225		
QY	181	ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAAGCTAAATATGTGTCTTAAT	240		
DB	226	ACTGCAAAACCACTTATACATAAAAAATGAGTATGAAATAAAAAATTTTTTGTCTCTACA	285		
QY	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATATTTTTCCT	300		
DB	286	TTTAAATTACCTTTTATTTGAAGCAATGATTTTTTACATCAATTTTATAATTTTTTCCCC	345		
QY	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360		
DB	346	GATGCAAAACCTTGGCTATGAAGTTATTGAAAAACCTTAAAGAAATTTTATGCTTATATAAAA	405		
QY	361	TTTCAGCAAAATTTATTTCAATCAAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGCC	420		
DB	406	TACAAATGAAATTTATTTTCAATAAAAGAAATTAATTCGGGCGTCTATATGTGTGCAATTTGCT	465		
QY	421	ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATTCGATTTTTTATCAAAATGGGTCA	480		
DB	466	ATTGCAATAGGATATATAAACCATCTATTTATGTGGCAATCGATTTTTTAAGAGGAGATGTT	525		
QY	481	TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTTTAAAAAT	540		
DB	526	ATTTATCCITTTGAAGCTATGAGTACAAATATAAAAAACAATCTTCTCGGAATAAAA- --	582		
QY	541	GATCGCTCGCACTATATCGGACATAGTAAAAATAACAGATAAAAGCTTTAGAAATTTCTA	600		
DB	583	GATTTCAAAACCTTCAAAATTCATCTTAAGGAATACGATATAGAAGCAATTAATAATTTGTA	642		
QY	601	GAATAAACTTACAAAATAAAACTATATTGCTTATGTCCTTAATAGTCTTTTACGCAATTTT	660		
DB	643	AAATCAATATACAAAGTTAATATCTACGCAATTTGTGATGATTCATTTATTTTGGCAATTCAT	702		



Db	378966	AGTCAAGCGATATCATCTTATGGAATATGATTTAAATGCGACTTTATTTTACAAAAA	378907	NAME/KEY: misc_feature LOCATION: (36636)..(36636) OTHER INFORMATION: n equals a,t,c, or g
Qy	607	ACTTACAAAATAAACTATATGCTTATGCTCTATATAGTCTTTTAGCAAAATTTATAGAA	666	FEATURE: NAME/KEY: misc_feature LOCATION: (40808)..(40810) OTHER INFORMATION: n equals a,t,c, or g
Db	378906	CATTATGGAGTAATAATTTATTTGCAATTCGCCAGAAAGTCTCTATGTAATTTATTTCT	378847	FEATURE: NAME/KEY: misc_feature LOCATION: (44416)..(44416) OTHER INFORMATION: n equals a,t,c, or g
Qy	667	CTAGCGCCAAATTTAAATTC A---AATTTTATCATACAAGAAAAAAATAACTACACTAAA	723	FEATURE: NAME/KEY: misc_feature LOCATION: (44905)..(44905) OTHER INFORMATION: n equals a,t,c, or g
Db	378846	TTATCACCAC TGAATAAACCAATTTACTTTTATCTCGAAGAAAAGAAAAATTACACAA	378787	FEATURE: NAME/KEY: misc_feature LOCATION: (44975)..(44975) OTHER INFORMATION: n equals a,t,c, or g
Qy	724	GATATACTCATACCTTCTAGTAGGGCTTATGGAATAATTTTCAAAAAATTTAATTTTAA	783	FEATURE: NAME/KEY: misc_feature LOCATION: (45593)..(45593) OTHER INFORMATION: n equals a,t,c, or g
Db	378786	GATATTTTAATTC-----GCCGAAGTTTGATATATAAAAAAATTTGGTATATATTC	378736	FEATURE: NAME/KEY: misc_feature LOCATION: (45732)..(45732) OTHER INFORMATION: n equals a,t,c, or g
Qy	784	AAAAATAAAATTAAGAAATGTTTATTACAAGTTTGATAAAAGATCTATTAAAGATTACCT	843	FEATURE: NAME/KEY: misc_feature LOCATION: (47036)..(47036) OTHER INFORMATION: n equals a,t,c, or g
Db	378735	AAACCAAGAATTTACAAAATCTGATTTTTCGTTGATCTGGATATATTAGTTTACCT	378676	FEATURE: NAME/KEY: misc_feature LOCATION: (51334)..(51334) OTHER INFORMATION: n equals a,t,c, or g
Qy	844	AGTGATATAAGCATTTATTTCAAAGGAAAAATAA	876	FEATURE: NAME/KEY: misc_feature LOCATION: (51602)..(51602) OTHER INFORMATION: n equals a,t,c, or g
Db	378675	AATGATATAAAACACGCTTTAAATCAAGAAAA	378643	FEATURE: NAME/KEY: misc_feature LOCATION: (51786)..(51786) OTHER INFORMATION: n equals a,t,c, or g
RESULT 52				
US-10-158-865-1/c				
Sequence 1, Application US/10158865				
Patent No. 6846651				
GENERAL INFORMATION:				
APPLICANT: Fleischmann et al.				
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag				
Patent No. 6846651				
TITLE OF INVENTION: Thereof, and Uses Thereof				
FILE REFERENCE: PB186P2C1D1				
CURRENT APPLICATION NUMBER: US/10/158,865				
CURRENT FILING DATE: 2002-06-03				
PRIOR APPLICATION NUMBER: US 09/557,884				
PRIOR FILING DATE: 2000-04-25				
PRIOR APPLICATION NUMBER: US 08/476,102				
PRIOR FILING DATE: 1995-06-07				
PRIOR APPLICATION NUMBER: US 08/426,787				
PRIOR FILING DATE: 1995-04-21				
NUMBER OF SEQ ID NOS: 1				
SOFTWARE: PatentIn version 3.1				
SEQ ID NO 1				
LENGTH: 1830121				
TYPE: DNA				
ORGANISM: Haemophilus influenzae				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (4747)..(4747)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (9921)..(9921)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (10150)..(10150)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (29298)..(29298)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (36543)..(36543)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (36551)..(36551)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (102696)..(102696)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (100091)..(100091)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (80024)..(80024)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (65313)..(65313)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (65309)..(65309)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (55369)..(55369)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (51805)..(51805)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (51602)..(51602)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (51334)..(51334)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (47036)..(47036)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (45732)..(45732)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (45593)..(45593)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (44975)..(44975)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (44905)..(44905)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (44416)..(44416)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (40808)..(40810)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (36636)..(36636)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (100091)..(100091)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (102696)..(102696)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (80024)..(80024)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (65313)..(65313)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (65309)..(65309)				
OTHER INFORMATION: n equals a,t,c, or g				
FEATURE:				
NAME/KEY: misc_feature				
LOCATION: (55369)..(55369)				
OTHER INFORMATION: n				

```
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (150841)..(150841)

; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature

Query Match 32.7%; Score 286.6; DB 3; Length 1830121;
Best Local Similarity 60.6%; Pred. No. 3.6e-46;
Matches 529; Conservative 0; Mismatches 329; Indels 15; Gaps 3;

Qy 10 GTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTAATTCAGGCTACCAAT 69
Db 379506 GTCATTATTCAGGTAATGGAACAAAGTTTAAATCAATTGACTAGTTATTACCTAAA 379447

Qy 70 GATTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTACTATCTCGTAA 129
Db 379446 GATTATGATGTTTCCGTTGCAATCAATTTTATTTTGAAGATCAATTTTCTTGGCAAG 379387

Qy 130 AAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACATACATTTTAA 189
Db 379386 AAAATAAAAAAGTATTTTAAATGTTCTCAATTTTIGAGCAATACATACATTTTATG 379327

Qy 190 CATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTTAATTAACACAA 249
Db 379326 CAATTAATTAATAATATGATGATGCTGATTAATTTTATCATCTTTTGTCAAT 379267

Qy 250 GCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCGTGATGTCAT 309
Db 379266 TTAGGAGATTTCAGAAATTAAGAAAAATTAATAATGTACAAAAATTAACAACAAGTTGAT 379207

Qy 310 TTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAATTTTCAAGAA 369
Db 379206 ATTGGACATTTATTTTAAACAAGCTACCCGCCCTTTGATGCTATTTTACAAATAACGAA 379147

Qy 370 ATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCCATAGCCCTA 429
Db 379146 TTATATGAAAATAGAGAAATTAATCAGGCGTTATATGTGTGAGTAGCCATGTAATG 379087

Qy 430 GGATACAAAGAAATTTATCTTTCGGGAATTCATTTTATC---AAAAATGGGTCTCTTAT 486
Db 379086 GGTATATAAGATCTTTTATTTAAACAGGTATTGATTTTATCAAGAAAAAGGGAATCCTTAC 379027

Qy 487 GCTTTGATACCAACAAAGAAAATCTTTTAAACCTAGCCCTGATTTTAAATAATGATCGC 546
Db 379026 GCATTTTCATCATCAAAAAGAAAAATATTATTAATTTATACCTTCTTTTTCACAAAATAA 378967

Qy 547 TCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAA 606
Db 378966 AGTCAAGCGATATCCCAATCTATGGAATATGATTTTAAATGCACTTTATTTTACAAAA 378907

Qy 607 ACTTACAAAAATAAACTATATTTGCTTATGCTCTCTATAGTCTTTTACGAAAATTTATAGAA 666
Db 378906 CATTATCGAGTAATATTTTATTTGCAATTTTCGCCAGAAAGCTCTCTATGTAATTTTCTCT 378847

Qy 667 CTAGCGCCAAATTTAAATTC---AATTTTATCATACAGAAAAATACTACACTAAA 723
Db 378846 TTATCACCACTGAATAACCCCAATTTACTTTTATTTCTCGAAGAAAAAGAAAAATTACACAA 378787

Qy 724 GATATACTCATCCTTCTAGTGAGGCTTATGGAAAATTTTCAAAAAATTTTAAATTTAAA 783
Db 378786 GATATTTTAAATTC-----GCCGAAAGTTTGTATATAAAAAATTTGGTATATATTC 378736

Qy 784 AAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAAAAGATCTATTAAGATTACCT 843
Db 378735 AAACCAAGAATTTACCAAAATCTGATTTTTCGGTTGATCTGGGATATATTACGTTTACCT 378676

Qy 844 AGTGATATAAGCATTTATTTCAAAAGGAAAAATAA 876
Db 378675 AATGATATAAACACACGCTTTAAAAATCAAGAAAA 378643
```





```
QY 644 GTCTTTTACGCAAAATTT-TATAGAACTAGCGCCAAATTTTAAATTCCAAATTTTATCATACAA 702
Db 200932 AATAATATATATATTTTAAATATAAATATATATAAATAATATATATTTTAAATATAAATA 200991
QY 703 GAAAAAATAACTACACATAAGATATATCTCATCTCTAGTGAGGCTTATGGGAAATTT 762
Db 200992 TAAAAATATATATTTTAAATATAAATAATATATAAATAATATATATTTTAAATATAAATA 201051
QY 763 TCAAAAAATATTAATTTTAAAAAATAAATAATTAAGAAAAATGTTTATTACAGTTTGATA 822
Db 201052 TATAAATAATATATATTTTAAATATAAATAATATATAAATAATATATTTA-TATATTTTAAAT 201110
QY 823 AAAGATCTATTAAAGATT 839
Db 201111 ATAAATATATTAATATT 201127

RESULT 56
US-09-949-016-15853
; Sequence 15853, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15853
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15853

Query Match 9.0%; Score 78.6; DB 3; Length 205044;
Best Local Similarity 47.7%; Pred. No. 3.3e-06;
Matches 323; Conservative 0; Mismatches 349; Indels 5; Gaps 3;

QY 164 TCTTTGACCAATACACACTTTAAACCAATTTAAATCCAAATCAAGAATATGAGACCGAAC 223
Db 200455 TATTTAAATTTATACATTTATATAGTTATATCATCAAAATTTTATCAATATATTTAA 200514
QY 224 TAATATGCTGTTCTTAATTAACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTT 283
Db 200515 TATAAATATATTTATATATATTTTAAATATAAATAATATAAATAATATATTTTATATATATTT 200574
QY 284 TTTACGATTAATTTTCCGTGCTCATTTGGGATATGATTTTTTAAACAACCTTTAAAGAT 343
Db 200575 TAAATATAAATATAAATAATATATTTTATATATATTTTAAATATAAATAATATAAATA 200634
QY 344 TTAATGCTTAATTTTAAATTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCT 403
Db 200635 TTTATATATATTTTAAATATAAATAATATAAATAATATATTTTATATATTTTAAATATAAAT 200694
QY 404 ATATGCTGTCAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTGATT 463
Db 200695 ATATAAATAATATTTTATATATATTTTAAATATAAATAATATAAATAATATATTTTATATATTT 200754
QY 464 TTTATCAAAATGGGTCACTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAG 523
Db 200755 TTAAATATAAATAATATAAATAATATATTTTATATATATTTTAAATATAAATAATATAAATA- 200813
QY 524 CCCCTGATTTTAAAAATGATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAA 583
Db 200814 --TATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATAA 200871
QY 584 AAGCTTTTAGAATTTCTAGAAAAAACTTACAAAAATAAACTATATATTGCTTTATGCTCTAATA 643
Db 200872 AATATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATAA 200931
QY 644 GTCTTTTACGCAAAATTT-TATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAA 702
Db 200932 AAAAAATATATATTTTAAATATAAATAATATAAATAATATATTTTAAATATAAATAATAA 200991
QY 703 GAAAAAATAACTACACATAAGATATATCTCATCTCTAGTGAGGCTTATGGGAAATTT 762
Db 200992 TAAAAATATATATTTTAAATATAAATAATATAAATAATATATTTTAAATATAAATAATAA 201051
QY 763 GAAAAAATAACTACACATAAGATATATCTCATCTCTAGTGAGGCTTATGGGAAATTT 762
Db 200992 TAAAAATATATATTTTAAATATAAATAATATAAATAATATATTTTAAATATAAATAATAA 201051
QY 763 TCAAAAAATATTAATTTTAAAAAATAAATAATTAAGAAAAATGTTTATTACAGTTTGATA 822
Db 201052 TATAAATAATATATATTTTAAATATAAATAATATATAAATAATATATTTA-TATATTTTAAAT 201110
QY 823 AAAGATCTATTAAAGATT 839
Db 201111 ATAAATATATTAATATT 201127

RESULT 56
US-09-949-016-12387
; Sequence 12387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12387
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12387

Query Match 9.0%; Score 78.6; DB 3; Length 223471;
Best Local Similarity 47.7%; Pred. No. 3.3e-06;
Matches 323; Conservative 0; Mismatches 349; Indels 5; Gaps 3;

QY 164 TCTTTGACCAATACACACTTTAAACCAATTTTAAATCCAAATCAAGAATATGAGACCGAAC 223
Db 170882 TATTTAAATTTATTAATATATTTTAAATATAAATAATATAAATAATATATTTTATCAATATATTTAA 170941
QY 224 TAATATGCTGTTCTTAATTAACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTT 283
Db 170942 TATAAATATATTTATATATATTTTAAATATAAATAATATAAATAATATATTTTATATATATTT 171001
QY 284 TTTACGATTAATTTTCCGTGCTCATTTGGGATATGATTTTTTAAACAACCTTTAAAGAAAT 343
Db 171002 TAAATATAAATATAAATAATATATTTTATATATATTTTAAATATAAATAATATAAATAATA 171061
QY 344 TTAATGCTTAATTTTAAATTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCT 403
```

		171062	T T T A T A T A T A T T T T P A A A T A T A A A A T A T A T T T A T A T A T T T T A A A T A T A A A T	171121
D <sub>b</sub>				
		404	A T A T G T C G A G T A G C C A T A G C C T A G G A T A C A A A G A A A T T A T C T T C G G A A T T G A T T	463
Q <sub>y</sub>				
		171122	A T A T A A A A T A T T T A T A T A T A T T T T A A R A T A T A A A A T A T A T T T A T A T A T A T T	171181
D <sub>b</sub>				
		464	T T T A T C A A A A T G G G T C A T C T T A T G C T T T T T G A T A C C A A C A G A A A A A T C T T T T A A A C T A G	523
Q <sub>y</sub>				
		171182	T T A A A T A A A A T A T A T A A A A T A T A T T T A T A T A T A T T T T A A A T A T A A A T A T A A A A A -	171240
D <sub>b</sub>				
		524	C C C C T G A T T T T A A A A A T G A T C G C T C G C A C T A T C G G A C A T A G T A A A A A T C A G A T A T A A	583
Q <sub>y</sub>				
	- - T A T A T A T T T T A A A T A T A A A T A T A T A T A T A T T T T T A A A T A T A A A T A T A T A A	171241	--TATATA TT TTTAAATATA AAATATATA TA TATTTTTTAATATAA ATATATAA	171298
D <sub>b</sub>				
		584	A A G C T T T A G A A T T T C T A G A A A A A A C T T A C A A A A T A A A A C T A T A T T G C T T A T C T C T A T A	643
Q <sub>y</sub>				
		171299	A A T A T A T A T T T T A A A T A T A A A A T A T A A A A T A T A T A T A T T T T A A A T A T A A A T A T A	171358
D <sub>b</sub>				
		644	G T C T T T T A G C A A T T T - T A T A G A A C T A G C G C A A A T T T A A A T T C A A A T T T A T C A T A C A A	702
Q <sub>y</sub>				
		171359	A A A A T A T A T A T T T T T A A A T A A A A T A T A T A A A A T A T A T A T A T T T T A A A T A A A A T A T A	171418
D <sub>b</sub>				
		703	G A A A A A A A T A A C T A C A C T F A A A G A T A T A C T C A T A C C T C T A G T G A G C C T T A T G G A A A A T T	762
Q <sub>y</sub>				
		171419	T A A A A A T A T A T A T T T T T A A A T A T A A A T A T A T A T A A A A A T A T A T A T A T T T T A A A T A T A A A T A	171478
D <sub>b</sub>				
		763	T C A A A A A A T A T T T T T A A A A A A T A A A A A T T A A A G A A A A T G T T T T A T T A C A A A G T T G A T A	822
Q <sub>y</sub>				
		171479	T A T A A A A A T A T A T A T T T T A A A T A A A A T A T A T A A A A A T A T A T A T A T T A T A T T T T A A A T	171537
D <sub>b</sub>				
		823	A A A G A T C T A T A A G A T T	839
Q <sub>y</sub>				
		171538	A T A A A T A T A T A A T A T T	171554
D <sub>b</sub>				

```

RESULT 57
US-09-949-016-12724
; Sequence 12724, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12724
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724

```

QY	224	TAATTTAGTGTCTTAAATTACAAACGAGCTCATCTAGAAAAATGAAAAATTTTCTGAAAAACCTT	283
DB	170942	TATAAAATATATTTTATATATATTTTAAAAATATAAAATATATAAAATATATTTTATATATATTT	171001
QY	284	TTTACGATTTATTTTCCCTGATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAAT	343
DB	171002	TAAATATAAATATATAAAATATATATTTTATATATTTTAAATATAAATATATAAATATA	171061
QY	344	TTAATGCTTATTTTAAAAATTTTACGAAATTTTATTTCAATCAAGAATATACCTCAGGGGTCT	403
DB	171062	TTTATATATATTTTAAAAATATAAAATATATAAAATATATTTATATATATTTTAAATATAAAT	171121
QY	404	ATATGCTGCGAGTACGACCTAGGATACAAAGRAATTTATCTCTCGGGAATTGATT	463
DB	171122	ATATAAAATATATTTTATATATATTTTAAATATAAATATATAAATATATTTTATATATATTT	171181
QY	464	TTTATCAAAATGGGTCTCTTATGCTTTTGTATGCCAAACAAGAAAACTCTTTTAAACCTAG	523
DB	171182	TTAAATATAAATATATAAAATATATTTTATATATATTTTAAATATAAATATATAAAATA-	171240
QY	524	CCCCTGATTTTAAAAATGATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAA	583
DB	171241	--TATATATTTTAAATATAAATATATAAATAATATATATTTTAAATATAAATATATAAA	171298
QY	584	AAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACTATATATGCTTGTCTTAATA	643
DB	171299	AAATATATATTTTAAATATAAATATATAAANAATATATATATTTTAAATATAAATATATA	171358
QY	644	GTCTTTTAGCAAAATTT--TATAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATACAA	702
DB	171359	AAAATATATATTTTAAATATAAATAATATAAAAAATATATATATTTTAAATATAAATATA	171418
QY	703	GAAAAAATAACTACACTAAAGATATATCTATACCTTCTAGCTAGGCTTATGAAAAATTT	762
DB	171419	TAAAAATATATATATTTTAAATATAAATATATAAANAATATATAAATATAAATATAAATA	171478
QY	763	TCAAAAAATATTTAAATTTTAAAAAATAAAAAATTAAGAAAAATGTTTATACAAAGTTGATA	822
DB	171479	TATAAAAAATATATATTTTAAATATAAATATAAATAATATAAANAATATATTTA-TATATTTTAAAT	171537
QY	823	AAAGATCTTATTAAGATT	839
DB	171538	ATAAATATATTAATATTT	171554

```

RESULT 58
US-09-949-016-12725
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12725
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725

```

Query Match 9.0%; Score 78.6; DB 3; Length 223471;  
Best Local Similarity 47.7%; Pred. No. 3.3e-06;  
Matches 323; Conservative 0; Mismatches 349; Indels 5; Gaps 3;  
QY 164 TCTTGAGCAATACTACACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAAC 223  
Db 170882 TAUTTTAAATTTATTATATATATAGTTTATATCATCAAAATTTTATCAATATATTTAAA 170941  
QY 224 TAATTATGCTTCTTAATTACCAACCAAGCTCATCTAGAAATGAATAATTTTGTAAAACTT 283  
Db 170942 TATAAATATATTTATATATATTTTAAATATAAATATAAATAATATATTTTATATATATTT 171001  
QY 284 TTTACGATTTATTTTCCGTGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAAT 343  
Db 171002 TAAATATAAATATAAATAATATATTTTATATATATTTTAAATATAAATATAAATAATA 171061  
QY 344 TTAATGCTTTATTTTAAATTTTCCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCT 403  
Db 171062 TTTATATATATTTTAAATATAAATATAAATAATATATTTTATATATTTTAAATATAAAT 171121  
QY 404 ATATGTCAGTAGCAGCATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGGAATTCGATT 463  
Db 171122 ATATAAAATATTTTATATATATTTTAAATATAAATATAAATAATATATTTTATATATAT 171181  
QY 464 TTTATCAAAATGGTCATCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAG 523  
Db 171182 TTAATATAAATATAAATAATATATTTTATATATATTTTAAATATAAATATAAATAATA - 171240  
QY 524 CCCTGATTTTAAAAATGATCGCTCGCATATATATCGGACATAGTAAATAATACAGATATA 583  
Db 171241 --TATATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATATAAATA 171298  
QY 584 AAGCTTTAGAAATTTCTAGAAAAAATCTTCAAAAATATAATTTATGCTTTATGCTCTAATA 643  
Db 171299 AATATATATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATATAATA 171358  
QY 644 GTCTTTAGCAAAATTT--TATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAA 702  
Db 171359 AAAATATATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATATAATA 171418  
QY 703 GAAAAAATAACTACACTAAGATATATCTACCTTCTAGTGGCTTATGGAAAAATTT 762  
Db 171419 TAAAAATATATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATAATA 171478  
QY 763 TCAAAAAATATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 822  
Db 171479 TATAAAAAATATATATTTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATA 171537  
QY 823 AAGATCTATTAAAGATT 839  
Db 171538 ATAATATATATAATTT 171554

RESULT 59  
US-09-949-016-12776  
; Sequence 12776, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIORITY FILING DATE: 2000-04-14  
; PRIORITY FILING DATE: 2000-10-20  
; PRIORITY FILING DATE: 2000-10-20  
; PRIORITY FILING DATE: 2000-10-03  
; PRIORITY FILING DATE: 2000-10-03  
; PRIORITY FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12776  
; LENGTH: 187169  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(187169)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12776

Query Match 7.9%; Score 69; DB 3; Length 187169;  
Best Local Similarity 46.5%; Pred. No. 0.00023;  
Matches 384; Conservative 0; Mismatches 415; Indels 26; Gaps 4;  
QY 24 AAATGGGCAAGTTTAAAGAAATTTGATTTTCAAGGCTACCAATGATTTTGTGATGATT 83  
Db 555 ATATATAAATATATATTTATATAGTATATAAATAATATAATATAATATAATATAATAA 614  
QY 84 TAGATGTAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAAATGCAAGCAGT 143  
Db 615 TATATAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 674  
QY 144 GTTTTACACCCCTAATTTCTCTTTGGACAATACACACTTTTAAACAAATTTTAAATCCAAA 203  
Db 675 ATATAAATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 734  
QY 204 TCAAGAAATATGAGACCGAACTA--ATTATGTTCTTAATTTCAACCAAGCTCATCTAGA 260  
Db 735 TATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 794  
QY 261 AAATGAAAAATTTCTGAAAAAATTTTACGATTTTCTCTGATGCTCATTTGGGATATGA 320  
Db 795 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 854  
QY 321 TTTTAAAAAACAATTTAAAGAAATTTAAATGCTTATTTTAAAT-----TTCAGCAAAATTTA 374  
Db 855 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 914  
QY 375 TTTCAATCAAGAAATTTACCTCAGGGCTTATATGTCGAGTAGCCATAGCCCTAGGATA 434  
Db 915 TATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 974  
QY 435 CAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGTCTATCTTTATGCTTTTGA 494  
Db 975 TAAATATATGATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1034  
QY 495 TACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAATGATCGCTCCGACATA 554  
Db 1035 TATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1094  
QY 555 TATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTTACAA 614  
Db 1095 TATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1153  
QY 615 AATTAACATATATGCTTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 674  
Db 1154 TAAATATATAATACATATATATTTATATATATATATATATATATATATATATATATATATA 1213  
QY 675 AAATTTAAATTTCAAAATTTTATACACAGAAAAAATAAATAAATAAATAAATAAATAAATAA 734  
Db 1214 AATATAATATATATATATTTTATATATATATATATATATAA-----ATATATAT 1257  
QY 735 ACCTTCTAGTGGCTTTATGGAATAATTTTCAAAAAAATTTTAAATTTTAAATAAATAAATAA 794  
Db 1258 TTTATATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1317  
QY 795 TAAAGAAATATGTTTATACAAAGTTGATAAAGATCTATTAAGATT 839  
Db 1318 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1362

RESULT 60  
US-09-949-016-15940

```
; Sequence 15940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15940
; LENGTH: 191569
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(191569)
; OTHER INFORMATION: n = A, T, C or G
; US-09-949-016-15940

Query Match      7.9%; Score 69; DB 3; Length 191569;
Best Local Similarity 46.5%; Pred. No. 0.00023;
Matches 384; Conservative 0; Mismatches 415; Indels 26; Gaps 4;

QY 24 AATGACCAAGTTTAAAGAAATTTGATTTATTCACGGCTACCAAAATGATTTTGTGATTT 83
DB 555 ATATATAAAATATATATATATATATATATATATATATATATATATATATATATATA 614
QY 84 TAGATGTAATCAATTTTATTTTCAAGATAAATACTATCTTGGTAAAGAAATGCAAGCAGT 143
DB 615 TATATATAATATAATATAATAAAATATATATATATATATATATATATATATATATAAT 674
QY 144 GTTTTACACCCCTAAATTTCTTCTTGGACAAATACACACTTTTAAACATTTAAATCCAAA 203
DB 675 ATATAATATACATAAAATATATATAAATATATATAATATATAAATAATAAATAATATAA 734
QY 204 TCAGAAATGAGACCGAACTA---ATTATGTTCTTAATTACCAACCAAGCTCATCTAGA 260
DB 735 TATATATAAAATAATAAAATATATATATATATAAATAATAAATAATAATATATATATA 794
QY 261 AATGAAATTTTGTAAAGAACTTTTACGATTTATTTTCTGATGCTCATTTGGGATATGA 320
DB 795 ATATATAAAATATATAAAATATATAAATATATAAATATATAAATATATAAATATATAA 854
QY 321 TTTTAAACCAACTTAAAGAAATTTTAAATGCTTATTTTAAAT-----TTCACGAAATTTA 374
DB 855 AATATATAAAATATATAAATAATATAAATAATATAAATAATATAAATAATATAAATAA 914
QY 375 TTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGTCAGTAGCCATAGCCCTAGATA 434
DB 915 TATAATATATAAATATAAATAATAAATAATAAATAATAAATAATAAATAATAATGGATATA 974
QY 435 CAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCATCTTATGCTTTTGA 494
DB 975 TAAATATATGATATAAATAATATAAATAATATAAATAATAAATAATAAATAATAAATAA 1034
QY 495 TACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAATGATCGCTCGACTA 554
DB 1035 TATATAAATATATAATATATAATATAATATATATATATATATATATATATAATATAA 1094
QY 555 TATCGGACATAGTAAATAATACAGATATAAAGCTTTTAGAATTTCTAGAAAAAATCTTACA 614
DB 1095 TATATAATATAAATAATATA-ATATAAATATATAAATAATAAATAATAAATAATAAATA 1153
QY 615 AATAAACTATATGCTTATGCTCCTTAATAGTCTTTTAGCAAAATTTTATAGAACTAGGCC 674
```

---

```
Db 1154 TAAATATATATAACATATATATATATATATATATATATATATATAAATATATATATA 1213
QY 675 AAATTTAAATTCAAATTTTATCATACAGAGAAAAAATACTACACTAAAGATATACTCAT 734
DB 1214 AATATAATATATATATATATATATATATATATATATATATATATATATATATAT 1257
QY 735 ACCTTCTAGTGAGGCTTATGCAAAATTTTCAAAAAATATTAATTTTAAAAAATAAAAAA 794
DB 1258 TTTATATATATATATAAATAATATATATATATATATATATATATATATATATATAT 1317
QY 795 TAAAGAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATT 839
DB 1318 AAAATAAATATATAAATAATATATATATAAATAAATAATATATATATAT 1362

RESULT 61
US-09-949-002-777/c
; Sequence 777, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 777
; LENGTH: 25590
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(25590)
; OTHER INFORMATION: n = A, T, C or G
; US-09-949-002-777

Query Match      7.9%; Score 68.8; DB 3; Length 25590;
Best Local Similarity 45.6%; Pred. No. 0.0002;
Matches 314; Conservative 0; Mismatches 372; Indels 2; Gaps 2;

QY 174 ATACTACACTTTAAACATTTAATCCAAATTCAGAAATATGAGACCGCAACTAATATATG 233
DB 4806 ATAAATATTTAAACCAATATATATATATATATATATATATATATATATATATATATA 4747
QY 234 TTCTAATTACACCAAGCTCATCTAGAAATGAAATTTTGTAAAGAACTTTTACAGATTA 293
DB 4746 ATATCATATATAATATATATATATATATATATATATATATATATATATATATATAA 4687
QY 294 TTTTCTGATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTA 353
DB 4686 TATATAATATATCTCATATATATATATATATATATATATATATATATATATATATATA 4627
QY 354 TTTTAAATTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGC 413
DB 4626 TATTAAATAATATAATATATATCTCATATATAATATATATATATATATATATATATATATC 4567
QY 414 AGTAGCCATAGCCCTCAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAA 473
DB 4566 ATATGTAATATTTAAATAATATAATATATATATATATATATATATATATATATATAA 4507
QY 474 TGGGTGATCTTATGCTTTTGTATACCAACAGAGAAATCTTTTAAAGAACTAGCCCTGATT 533
DB 4506 TATCATATGTAATATATAATATAATATATATATATATATATATATATATATATATATAA 4447
QY 534 TAAAAATGATCGCTCGCACTATATCGGACACTAGTAAAAAATACAGATATAAAGCTTTAGA 593
DB 4446 TATATATCATATGTAATATATATATATATATATATATATATATATATATATATATAA 4387
```



Db 204 ATTTTAAATATAAATATAAATAATATATATATTTTAAATATAAATATAAATAATATAT 145  
QY 683 ATTCAAATTTTATCATACAGAGAAAAATAACTACACTAAGATATACTCATACCTTCTA 742  
Db 144 ATATTTTAAATATAAATATAAATAATATATATATTTTAAATATAAATATAAATAATAT 85  
QY 743 GTGAGCCTTATGGAATTTTCAAAAAATTTAAATTTTAAAAAATAAATAATTTAAAGAA 802  
Db 84 ATATATTTTAAATATAAATATAAATAATATATATATTTTAAATATAAATATAAATAAT 25  
QY 803 ATGTTTATTACAAAGTTGATAAA 824  
Db 24 ATATATATTTTAAATATAAATA 3

## RESULT 64

US-09-949-016-37164/c  
; Sequence 37164, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37164  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-37164

Query Match 7.6%; Score 66.4; DB 3; Length 601;

Best Local Similarity 49.6%; Pred. No. 0.00041;

Matches 279; Conservative 0; Mismatches 276; Indels 7; Gaps 4;

QY 267 AAATTTTGAAAAACCTTTTACGATTTATTTTCTCGATGCTCATTTGGGATATGATTTTT 326  
Db 561 AAATTTTATCAATATATTTTAAATATAAATATATTTATATATATTTTAAATATAAATAT 502  
QY 327 TAAA--CAACTTAAAGAAATTTAATGCTTATTTTAAATTTTCAAGAAATTTTCAATCAA 384  
Db 501 AAAAAATATATTTATATATATTTTAAATATAAATATAAATAATATATTTTATATATTTT 442  
QY 385 AGAATTTACCTCAGGGTCTATATGTGTGAGTAGCCATAGCCCTAGGATACAAAGAAAT 444  
Db 441 AAATATAAATATAAATAATATTTTATATATATTTTAAATATAAATATAAATAATATAT 382  
QY 445 TATCTTTTCGGGAATTTGATTTTATCAAAATGGTCTATTTATGCTTTTGGATACCAACAA 504  
Db 381 TATATATATTTTAAATATAAATAATATATATTTTAAATATAAATATAAATAATATAT 322  
QY 505 GAAATCTTTTAAACTAGCCCTGATTTTAAAAATGATCGCTCGCAGCTATATCGGACAT 564  
Db 321 ATAAAAATATTTTATATATATTTTAAATATAAATATAAATAATATATTTTATAT--ATAT 265  
QY 565 AGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACT 623  
Db 264 TTTAAATATAAATATAAATAATATATATTTTAAATATAAATAATATAAATAATATAT 205  
QY 624 ATATTGCTTTATGCTCTTAATAGTCTTTTAGCAAAATTTT-ATAGAACTAGCGCCAAATTTAA 682  
Db 204 ATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATAATAT 145  
QY 683 ATTCAAATTTTATCATACAGAGAAAAATAACTACACTAAGATATACTCATACCTTCTA 742

Db 144 ATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATATAAATAATAT 85  
QY 743 GTGAGCCTTATGGAATTTTCAAAAAATTTAAATTTTAAAAAATAAATAATTTAAAGAA 802  
Db 84 ATATATTTTAAATATAAATATAAATAATATATATATTTTAAATATAAATAATATAAATAAT 25  
QY 803 ATGTTTATTACAAAGTTGATAAA 824  
Db 24 ATATATATTTTAAATATAAATA 3

## RESULT 65

US-09-949-016-145868/c  
; Sequence 145868, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 145868  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-145868

Query Match 7.6%; Score 66.4; DB 3; Length 601;

Best Local Similarity 49.8%; Pred. No. 0.00041;

Matches 279; Conservative 0; Mismatches 276; Indels 7; Gaps 4;

QY 267 AAATTTTGAAAAACCTTTTACGATTTATTTTCTCGATGCTCATTTGGGATATGATTTTT 326  
Db 561 AAATTTTATCAATATATTTTAAATATAAATATATTTATATATATTTTAAATATAAATAT 502  
QY 327 TAAA--CAACTTAAAGAAATTTAATGCTTATTTTAAATTTTCAAGAAATTTTCAATCAA 384  
Db 501 AAAAAATATATTTATATATATTTTAAATATAAATATAAATAATATATTTTATATATTTT 442  
QY 385 AGAATTTACCTCAGGGTCTATATGTGTGAGTAGCCATAGCCCTAGGATACAAAGAAAT 444  
Db 441 AAATATAAATATAAATAATATTTTATATATATTTTAAATATAAATATAAATAATATAT 382  
QY 445 TATCTTTTCGGGAATTTGATTTTATCAAAATGGTCTATTTATGCTTTTGGATACCAACAA 504  
Db 381 TATATATATTTTAAATATAAATAATATATATTTTAAATATAAATATAAATAATATAT 322  
QY 505 GAAATCTTTTAAACTAGCCCTGATTTTAAAAATGATCGCTCGCAGCTATATCGGACAT 564  
Db 321 ATAAAAATATTTTATATATATTTTAAATATAAATAATATAAATAATATATTTTATAT--ATAT 265  
QY 565 AGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACT 623  
Db 264 TTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATAATATAT 205  
QY 624 ATATTGCTTTATGCTCTTAATAGTCTTTTAGCAAAATTTT-ATAGAACTAGCGCCAAATTTAA 682  
Db 204 ATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATAATAT 145  
QY 683 ATTCAAATTTTATCATACAGAGAAAAATAACTACACTAAGATATACTCATACCTTCTA 742  
Db 144 ATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAAATAATAT 85





Db	24	ATATATATTTTAAATATAATAA 3	604	SAACAMANAATTCYCTVAANNATYTWANNATGCMNATKTAATWNTNNNNNAGTWTNNNN 666
RESULT 68				
US-09-806-708B-23				
; Sequence 23, Application US/09806708B				
; Patent No. 6784342				
; GENERAL INFORMATION:				
; APPLICANT: The University of British Columbia				
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants				
; FILE REFERENCE: 4810-58741				
; CURRENT APPLICATION NUMBER: US/09/806,708B				
; CURRENT FILING DATE: 2001-04-03				
; PRIOR APPLICATION NUMBER: US 60/147,133				
; PRIOR FILING DATE: 1999-08-04				
; NUMBER OF SEQ ID NOS: 23				
; SOFTWARE: PatentIn version 3.0				
; SEQ ID NO 23				
; LENGTH: 1055				
; TYPE: DNA				
; ORGANISM: Artificial sequence				
; FEATURE:				
; NAME/KEY: promoter				
; LOCATION: (1)..(1055)				
; OTHER INFORMATION: consensus sequence of A.t. and L.a. PAB1 promoters				
US-09-806-708B-23				
Query Match 7.5%; Score 65.6; DB 3; Length 1055;				
Best Local Similarity 21.4%; Pred. No. 0.00061;				
Matches 180; Conservative 186; Mismatches 471; Indels 3; Gaps 1;				
QY	29	GACCAAGTTTAAAGAAATGATTATCAAGGTACCAATGATTGATGATTTAGAT 88		
Db	4	SAKWAARMYAKYAGVNNNTGRTTKGTWYTCANNNTGKCYARRWGWMTTAYYWTAT 63		
QY	89	GTAATCAATTTTATTTGAAGATAAATCTATCTTGGTAAATAATGCAAGCAGTGTITT 148		
Db	64	KWGTGAWRTWRWAAKTRKWCSTAMNNAMTWTCTARKWRTGRTWTKTNNATGTR 123		
QY	149	ACACCCTTAATTTCTTCTTGAGCAATACACTTAAACATTTTAAATTCACAAATCAAG 208		
Db	124	WNTGYMNNNGCSTWARRYKTRWVCYTAMYGASWAGNASTRRRTTYTWRKWKCR 183		
QY	209	AATATGAGACCGAAGCTAATATGTTCTTAATTACACCAAGCTCATAGAAATGAAA 268		
Db	184	KSARARATGRARYMRAWYTAWARRTGWTKAMAYAAWNNNNNNNNNNNNNNNNNNNN 243		
QY	269	ATTTTGTAATAAATTTTACGATTATTTTCTGATGCTCATTTGGGATATGATTTTTTA 328		
Db	244	KSNCTCTAGTTTTRKATCCWAYTCGAGWATKWKTKTSAAGMTWNNNNNNNNNTTTKA 303		
QY	329	AACAACCTTAAGAAATTAATGCTTATTTTAAATTTTCAAGAAATTTATTTCAATCAAGAA 388		
Db	304	ANYAARMWMSATTTWAAATSRKTWTYGRKTAMNNNGTTCWRMWAATWKWMT 363		
QY	389	TTACCTCAGGGCTATATGTGTGAGTAGCCATAGCCCTAGGATACAAAGAAATTTATC 448		
Db	364	KGTTTNNNGRGTGTGTTTKMAATTTTAKANNCTTAAWKTKTCTMNTTAAKATTYWATC 423		
QY	449	TTTC---GGGAATTTGATTTTATCAATGGGTCATCTTATGCTTTTATGATACCAACAAG 505		
Db	424	YKSMWNGTSYRYAARYTWAWWTRRYAYANNNTKTWKACTWTYKRCCTTANNNTAA 483		
QY	506	AAAACTTTTAAACATGAGCCCTGATTTTAAATATGATCGCTCGCACTATATCGGCATA 565		
Db	484	WYTKSSANCTSRTRWKTNCWRAGSKTASGRAYARAYWTGYKWTAWAYWCWWTYYRA 543		
QY	566	GTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACTAT 625		
Db	544	GAATWYMTSATCYCATAAATTAGTCAGAGGSGTAKGNNNNNNNNNNNNNNNNNNNN 603		
QY	626	ATTGCTTAGTCTCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTAAAT 685		
Query Match 7.5%; Score 65.4; DB 2; Length 6243;				
Best Local Similarity 49.9%; Pred. No. 0.0008;				
Matches 191; Conservative 0; Mismatches 191; Indels 1; Gaps 1;				
QY	486	TGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAATGATCG 545		
Db	1089	TTCCITTCGGAACCTAACTAAAGTAAAAAATTTATCTTTTACAACTCCCAAAAAAGAAC 114		
QY	546	CTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTAGAAA 605		

Db 1149 GGTACAAAGTACCCTATATACACAGCGCTAAAAAATGAGGGTAAAAAATAAAAAAATAAAA 1208  
Qy 606 AACTTACAAAATAAACTATATTCCTTATGTCCTAAATAGTCTTTTAGCAAAATTTTATAGA 665  
Db 1209 AAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAA 1268  
Qy 666 ACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAAGAAAAAATAAC-TACACTAAAG 724  
Db 1269 TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 1328  
Qy 725 ATATACATACACCTCTAGTGAGCGCTTATGGAAGATTTTCAAAGAAATTAATTTTAAAA 784  
Db 1329 AAAATATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 1388  
Qy 785 AAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 844  
Db 1389 TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 1448  
Qy 845 GTGATATAAAGCATTTATTTTCAAA 867  
Db 1449 ATCTTTGAAGAAAGATATAAA 1471

RESULT 70  
US-09-949-016-16110/c  
; Sequence 16110, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16110  
; LENGTH: 231129  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16110

Query Match 7.4%; Score 65.2; DB 3; Length 231129;  
Best Local Similarity 47.2%; Pred. No. 0.0013;  
Matches 265; Conservative 0; Mismatches 293; Indels 4; Gaps 2;  
Qy 318 TGATTTTTTAAACAACCTTAAAGAAATTTAAATGCTTTATTTTAAATTTTCAAGAAATTTATTT 377  
Db 72963 TGCCTTGTCTCAAGACCAAAATGCCATATATATAATATATATATATTTATATATATATATAT 72904  
Qy 378 CAATCAAGAATTAACCTCAGGGGCTATATATGTGCGAGTAGCCCTAGGATACAA 437  
Db 72903 AAATATATATATATATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATA 72844  
Qy 438 AGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCATCTTAT-GCTTTTGATA 496  
Db 72843 TATATATATATATATATTAAGTATATAATATATATAATATATATAATATATAATAATA 72784  
Qy 497 CCAACAGAGAAATCTTTTAAATCTAGCCCTGATTTTAAATGATCGCTCGCACTATA 556  
Db 72783 TAAATACATAATATATATTTACATATATAAATATATAATATATAATATATAATATATA 72724  
Qy 557 TCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAATTTACAAAA 616  
Db 72723 TATTATATATAATATAAATATATATAATATAATATAAATATAAATATAATATAATATA 72664  
Qy 617 TAAAACTATATGCTTATGCTCCTAAATAGCTTTTAGCAAAATTTTATAGAACTAGGCCAA 676

Db 72663 TATATAATATATATATATATAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA 72604  
Qy 677 ATTTAAATTCAAATTTTATCATACAGAAAAAATAAATCTACACTAAAGATATACTCATAC 736  
Db 72603 TATATATTTTATATATAAATAAATAATATATATATATATATAAATAAATAATATATAT 72544  
Qy 737 CTCTAGTGAGCGCTTATGGAAGATTTTCAAAGAAATAT--TAAATTTTAAAAAATAAAAA 793  
Db 72543 ATATAATATATATATATATATATATATATATAAATAAATAATATATATATATATAT 72484  
Qy 794 TTAAGAAAAATGTTTATTTATACAGTTGATAAAGATCTATTTAAGATTAACCTAGTGATATA 853  
Db 72483 TTATATAAATATAATAATATATATATATATATATATATATATATATATAAATAATATA 72424  
Qy 854 AGCATTATTTCAAAGGAAAAATA 875  
Db 72423 TATTATATATATATAAATAATA 72402

RESULT 71  
US-09-949-016-11934/c  
; Sequence 11934, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11934  
; LENGTH: 266293  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-11934

Query Match 7.4%; Score 65.2; DB 3; Length 266293;  
Best Local Similarity 47.2%; Pred. No. 0.0013;  
Matches 265; Conservative 0; Mismatches 293; Indels 4; Gaps 2;  
Qy 318 TGATTTTTTAAACAACCTTAAAGAAATTTAAATGCTTTATTTTAAATTTTCAAGAAATTTATTT 377  
Db 60127 TGCCTTGTCTCAAGACCAAAATGCCATATATATAATATATATATATATATATATATATAT 60068  
Qy 378 CAATCAAGAATTAACCTCAGGGGCTATATGTGCGAGTAGCCCTAGGATACAA 437  
Db 60067 AAATATATATATATATAAATAAATAAATAATATATATATATATATATATATAAATAATA 60008  
Qy 438 AGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCATCTTAT-GCTTTTGATA 496  
Db 60007 TATATATATATATATATTAAGTATATAATATATAATATAATATAATATAAATAATA 59948  
Qy 497 CCAACAAGAAAAATCTTTTAAACCTAGCCCTGATTTTAAAGATGATCGCTCGCACTATA 556  
Db 59947 TAAATACAAAAATATATATATACATATAAATAATATAAATAATATAAATAAATAATA 59888  
Qy 557 TCGGACATAGTAAAAATACAGATATAAAGCTTTTAAAGATTTCTAGAAAAAATTTACAAAA 616  
Db 59887 TATTATATATAATATAAATAATATATAATATAAATAATATAAATAATATAAATAATA 59828  
Qy 617 TAAAACTATATGCTTCTCCTAAATAGCTTTTAGCAAAATTTTATAGAACTAGGCCAA 676  
Db 59827 TATATAATATATATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 59768







Db 11641 ATATATATATATAAATATATATATATAAATATATATAATATATATAATATATATAATATATATA 11582  
Qy 614 AAATAAACTATATTGCT-TATGTCCTAAATAGTCTTTTAGCAAAATTTTATAGAACTAGCG 672  
Db 11581 AATATATAATATAAATAATAT 11522  
Qy 673 CCRAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATTAACACTACCTAAAGATATATCTC 732  
Db 11521 ATA 11462  
Qy 733 ATACCTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTCAAAAAATATAA 792  
Db 11461 TATAAATA 11402  
Qy 793 ATTAAGAAATGTTTATTACAGTTGTAAGATCTTATTAAGATTAACCTAGTAGATATA 852  
Db 11401 TATATATAATATATATATATATAAATAATATATATATATATATATATATATATATATATA 11342  
Qy 853 AAGCATTATTT 863  
Db 11341 TATAAATATAT 11331

RESULT 78  
US-09-662-254B-23  
; Sequence 23, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bowden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1XCI  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-662-254B-23

Query Match 7.4%; Score 64.4; DB 3; Length 50000;  
Best Local Similarity 46.8%; Pred. No. 0.0015;  
Matches 203; Conservative 0; Mismatches 231; Indels 0; Gaps 0;  
Qy 441 AATTATCTTCGGGAATTCGATTTTATCAAAATGGGTCACTTATGCTTTTGATACCAA 500  
Db 47317 AATTATTTTATTAATAATATATATATATATAAATCATCATCTACTAGTATATATAT 47376  
Qy 501 ACAAGAAAATCTTTTAAAACTAGCCCTGATTTTAAATAATGATCGCTCGCACTATATCGG 560  
Db 47377 AATAAAAAATGTAATAATTTTACAAAATATTTTAAATAATATATGATCAAGCAATAATA 47436  
Qy 561 ACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTAGAAAAAATTTACAAAAATAA 620  
Db 47437 CTGATTTCTTTTAAATGGAATAATATAAATGGGATAAAATTTAAATAATATAATAA 47496  
Qy 621 ACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTTATAGAACTTAGCGCCAAAT 680  
Db 47497 AATATTTATATAAATAATTTTGTATGATATAAATAACAGAAATTAATATGTTTATGAT 47556  
Qy 681 AAATTCAAAATTTTATCATACAAGAAAAAATAAATACTACATAAAGATATACCTCATACCTTC 740  
Db 47557 GACTGTAATTAATTTTAAATAAATAATTTATATAAATAATTTATATAAATAATTAATGA 47616  
Qy 741 TAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTCAAAAAATTTTCAAAAAATTAAGA 800

Db 47617 TATAGACTGTTTTTTTACTCTTCTTGTAGATATAAATATAAATAATGTATAATTTTACA 47676  
Qy 801 AAATGTTTTTATACAAAGTTGATAAAGAGCTTATTAAGATTACCTAGTGATATAAAGCATTA 860  
Db 47677 AATATTTTAAAAAAATTCAAAATATAAATTTTAAATTTATATATTCGATAAAGTTATA 47736  
Qy 861 TTTCAAAAGGAAAAAT 874  
Db 47737 TGGAAACCAACAT 47750

RESULT 79  
US-09-902-540-1357  
; Sequence 1357, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1357  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(612)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1357

Query Match 7.3%; Score 64.2; DB 3; Length 612;  
Best Local Similarity 47.8%; Pred. No. 0.0011;  
Matches 183; Conservative 0; Mismatches 200; Indels 0; Gaps 0;  
Qy 494 ATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAATAATGATCGCTCGCACT 553  
Db 207 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 266  
Qy 554 ATATCGGCACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTAGAAAAAATTTACA 613  
Db 267 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 326  
Qy 614 AAATAAACTATATGCTTATGCTCTTAATAGTCTTTTACCAATTTTATAGAACTAGCGC 673  
Db 327 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 386  
Qy 674 CAATTTTAAATTCAAATTTTATCATACAGAAAAATAAATAAATAAATAAATAAATAAATA 733  
Db 387 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 446  
Qy 734 TACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTCAAAAAAT 793  
Db 447 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 506  
Qy 794 TTAAGAAAAATGTTTATTTTCAAGTTGATAAAGAGCTTATTAAGATTACCTAGTGTATATA 853  
Db 507 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 566  
Qy 854 AGCATTATTTCAAGGAAAAATAA 876  
Db 567 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 589

RESULT 80  
US-09-806-708B-22/c



Db 8176 TCATGAATATAAATCATCTATATATAAATCAATATTATTTTAAAAATATCTTTAAATA 8117  
Qy 569 AAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAATCTTACAAAATAAAACTATATT 628  
Db 8116 TTAATATCTTTTATATTATATTATTTATATAATAAATTTAAATTTTAAAGTATAAT 8057  
Qy 629 GCTTATGCTCTATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTAAATTCAA 688  
Db 8056 AATGGTTTATATCAATAATATGTTGTTTAAATGATAATTTTATTTTATTTTAA 7997  
Qy 689 ATTTTATCAACAGAAAAAATACTACACTAAAGATATACTCATACCTTCTAGTGAGG 748  
Db 7996 CAITTT-----TAAAGAATTATATTATTTATATAAATTTATCTAAATTTAAATTCGATT 7943  
Qy 749 CTTATGAAAAATTTTCAAAAAATTTAAATTTTAAAAAAATTAATAATTAAGAAATGTTT 808  
Db 7942 AAATTTAAATTTATTTCTATAAAATTTATATTACTATATCAATATTATTAAAGTGAAT 7883  
Qy 809 ATTACAAGTTGATAAAAAATCTATTAAAGATTACCTAGTGATATAAAGCAATTTTCAAG 868  
Db 7882 ACTTCAATAGTATTAATAATTACTGATTTTTTCATATATCTATATGCATATTTTCRAAT 7823  
Qy 869 GAAAAATAA 876  
Db 7822 TTTAATGA 7815

RESULT 82  
US-09-949-016-15348  
; Sequence 15348, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15348  
; LENGTH: 134987  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15348

Query Match 7.3%; Score 63.6; DB 3; Length 134987;  
Best Local Similarity 44.5%; Pred No. 0.0024;  
Matches 294; Conservative 0; Mismatches 364; Indels 2; Gaps 1;  
Qy 177 CTACACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATTTAGTGTC 236  
Db 11230 CGACAGAGCTAGACTTCTCTCAAAAAAAGAAATCTCTTTGGTTTATAT 11289  
Qy 237 TAATTACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTT 296  
Db 11290 ATATTTTTTATATATAATATAATTAATAATATAATATAATATAATATAATAA 11349  
Qy 297 TCCTGATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTT 356  
Db 11350 TATATAAATATATATATATATATATTTTTTATATATATATATATATATATATAT 11409  
Qy 357 TAAATTTTCAAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGT 416  
Db 11410 TATATATTTAT 11469  
Qy 417 AGCCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCGATTTTATCAAAATGG 476

Db 11470 ATATTTATATAATAAT 11529  
Qy 477 GTCATCTATGCTTTTGATACCAACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAA 536  
Db 11530 ATATATTATATATTAT 11589  
Qy 537 --AAATGATCCTCGCACATATATCGGCATATATCGGCATATATCGGCATATATCGGCAT 594  
Db 11590 TTAT 11649  
Qy 595 TTCTAGAAAAAATCTTACAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTT 654  
Db 11650 ATATTTAT 11709  
Qy 655 AATTTTATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAC 714  
Db 11710 ATATATATAATATATATGTTATATATATATATATATATATATATATATATATATAT 11769  
Qy 715 TACACTAAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATT 774  
Db 11770 TAT 11829  
Qy 775 AATTTTAAAAAATAAAATTTAAAGAAATGTTTATTTTCAAGTTGATTAAGAACTATTTA 834  
Db 11830 TAATATATATAAAACATTA 11889

RESULT 83  
US-09-949-016-15349  
; Sequence 15349, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15349  
; LENGTH: 134987  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15349

Query Match 7.3%; Score 63.6; DB 3; Length 134987;  
Best Local Similarity 44.5%; Pred No. 0.0024;  
Matches 294; Conservative 0; Mismatches 364; Indels 2; Gaps 1;  
Qy 177 CTACACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATTTAGTGTC 236  
Db 11230 CGACAGAGCTAGACTTCTCTCAAAAAAAGAAATCTCTTTGGTTTATAT 11289  
Qy 237 TAATTACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTT 296  
Db 11290 ATATTTTTTATATATAATATAATTAATAATATAATATAATATAATATAATAATAA 11349  
Qy 297 TCCTGATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTT 356  
Db 11350 TATATAAATATATATATATATATATTTTTTATATATATATATATATATATATATAT 11409  
Qy 357 TAAATTTTCAAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGT 416  
Db 11410 TATATATTTAT 11469



```
QY 417 AGCCATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTCATTTTATCAAAATGG 476
Db 11470 ATATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 11529
QY 477 GTCATCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCCCTGATTTAA 536
Db 11530 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11589
QY 537 --AAATGATCGCTCGGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAA 594
Db 11590 TTATATATATATATATATATATATATATATATATATATATATATATATATATAT 11649
QY 595 TTCTAGAAAAAATCTACAAAAATAAACTATATTTGCTTATGCTCCTAATAGTCTTTTAGCA 654
Db 11650 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11709
QY 655 AATTTTATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAAAGAAAAATAAC 714
Db 11710 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11769
QY 715 TACACTAAAGATATATCTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATT 774
Db 11770 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11829
QY 775 AATTTTAAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTA 834
Db 11830 TAATATATATAAAACATATATAATATATATATATATATATATATATATATATATAT 11889
```

## RESULT 84

```
US-09-949-016-15350
; Sequence 15350, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15350
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15350
```

```
Query Match 7.3%; Score 63.6; DB 3; Length 134987;
Best Local Similarity 44.5%; Pred. No. 0.0024;
Matches 294; Conservative 0; Mismatches 364; Indels 2; Gaps 1;

QY 177 CTACACTTTAAACATTTAATCCAAATCAAGAATATGACACCGAAGTCTTTGGTTTATAT 236
Db 11230 CGACAGCTAGACTCTTCTCAAAAAAAGGAATCTCTTTGGTTTATAT 11289
QY 237 TAATTACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTT 296
Db 11290 ATATTTTTTATATATAATATATATATATATATATATATATATATATATATATATAT 11349
QY 297 TCCTGATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTT 356
Db 11350 TATATAAATATATATATATATATATTTTTTATATATATATATATATATATATATATAT 11409
QY 357 TAAATTTACGAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTCAGT 416
Db 11410 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11469
```

```
QY 417 AGCCATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTCATTTTATCAAAATGG 476
Db 11470 ATATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 11529
QY 477 GTCATCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCCCTGATTTAA 536
Db 11530 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11589
QY 537 --AAATGATCGCTCGGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAA 594
Db 11590 TTATATATATATATATATATATATATATATATATATATATATATATATATATAT 11649
QY 595 TTCTAGAAAAAATCTACAAAAATAAACTATATTTGCTTATGCTCCTAATAGTCTTTTAGCA 654
Db 11650 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11709
QY 655 AATTTTATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAAAGAAAAATAAC 714
Db 11710 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11769
QY 715 TACACTAAAGATATATCTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATT 774
Db 11770 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11829
QY 775 AATTTTAAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTA 834
Db 11830 TAATATATATAAAACATATATAATATATATATATATATATATATATATATATATAT 11889
```

## RESULT 85

```
US-09-949-016-15507
; Sequence 15507, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15507
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15507
```

```
Query Match 7.3%; Score 63.6; DB 3; Length 134987;
Best Local Similarity 44.5%; Pred. No. 0.0024;
Matches 294; Conservative 0; Mismatches 364; Indels 2; Gaps 1;

QY 177 CTACACTTTAAACATTTAATCCAAATCAAGAATATGACACCGAAGTCTTTGGTTTAT 236
Db 11230 CGACAGCTAGACTCTTCTCAAAAAAAGGAATCTCTTTGGTTTATAT 11289
QY 237 TAATTACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTT 296
Db 11290 ATATTTTTTATATATAATATATATATATATATATATATATATATATATATATATAT 11349
QY 297 TCCTGATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTT 356
Db 11350 TATATAAATATATATATATATATATTTTTTATATATATATATATATATATATATATAT 11409
QY 357 TAAATTTACGAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTCAGT 416
```

Db	11410	TATATATTTATATATTTTATATATATATATATATTTATATATATATATATATTTATATATATATATAT	11469
Qy	417	AGCCATAGCGCTAGGATACAAAGAAAATTTATCTCTTCGGGAAATTGATTTTATCAAAATGG	476
Db	11470	ATATTTATATATATATATATTTATATATATTTATATATATATATATATATATATATATTTAT	11529
Qy	477	GTCATCTTATGCTTTTGTATACCAACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAA	536
Db	11530	ATATATATATATATATATATATATATATATATATATATTTATATATATATATATATATATATATA	11589
Qy	537	--AAATGATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAA	594
Db	11590	TTATATATATATATTTATATATTTATATATATATATATTTTATATATATATATATATTTAT	11649
Qy	595	TTTCTAGAAAAAACTTACAAAAATAAACTATATTCGTTATGTCCTAAATAGTCTTTTAGCA	654
Db	11650	ATATTATATATTTATATTTATATATATATATATATATATATATATATATATATATATTT	11709
Qy	655	AATTTTATAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATACAAGAAAAAATAAC	714
Db	11710	ATATATATAATATTTGTATATATTTATATATATCTATAATATTTATATATATATATATATTA	11769
Qy	715	TACACTAAAGATATATCTCATGCTTCTAGTGAAGCTTATCGAAAAATTTTCAAAAAAATTT	774
Db	11770	TATATTATAATATATTTATATATTTATATATATATATTTTATATATAAATATGTAATAATA	11829
Qy	775	AATTTTAAAAAATAAAAAATTTAAAGAAAATGTTTTATTACAAGTTGATTAAGATCTATTATA	834
Db	11830	TAATATATATAAAACATATAAATATATATATTTATATATATTTATATATATATATATATATATTA	11889

RESULT 86

```

US-09-949-016-15508
; Sequence 15508, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN
; TITLE OF INVENTION: WITH HUMAN DISEASE, I
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15508
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15508

```

	Query Match	7.3%;	Score 63.6;	DB 3;	Length 134987;
	Best Local Similarity	44.5%;	Pred. No. 0.0024;		
	Matches 294;	Conservative 0;	Mismatches 364;	Indels 2;	Gaps 1;
Qy	177	CTACACTTTAAACATTTTAATCCAAANTCAAGAAATGAGACCGAACTAAATTATGTGTTTC	236		
Db	11230	CGACAGCTAGACTTCTCTCAAAAAAATAAAAAAGGAATCTCTTTGGTTTTATAT	11289		
Qy	237	TAATTACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTTACGATTATTTT	296		
Db	11290	ATATTTTTTTTATATATAATAATATATATATAATAATATATATATATATATAATAATAA	11349		
Qy	297	TCTGTATGCTCATTTGGGATATGAATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTT	356		
Db	11350	TATATAAAAT	11409		
Ov	357	TAAATTTTCAGAAATTTATTTTCAATCAAAGAAATTACCTCAGGGTCTATATGTGTGCAGT	416		

		TATATATTATATATTTTAT	11410
D <sub>b</sub>			11469
		AGCCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGG	476
Q <sub>y</sub>			
		ATATTTATATAAT	11470
D <sub>b</sub>			11529
		GTCATCTTATGCTTTTGATACCAAACAAGAAATCTTTTAAAACTAGCCCCGTGATTTTAA	536
Q <sub>y</sub>			
		ATATATTTAT	11530
D <sub>b</sub>			11589
		--AAATGATCGCTCCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAA	594
Q <sub>y</sub>			
		TTATATATTATATATTTATATATTTATATATATATATATATATATATATATATATATAT	11590
D <sub>b</sub>			11649
		TTTCTAGAAAAAACHTACAAAATAAAACATATTTGCCTTAGTCCTTAATAGTCTTTTAGCA	654
Q <sub>y</sub>			
		ATATTTATTT	11650
D <sub>b</sub>			11709
		AATTTTATPAGAACTAGCGCAAATTTAAATTCAAATTTTATCATACAAGAAAAAAATAAC	714
Q <sub>y</sub>			
		ATATATATAATATATTTGTATTA	11710
D <sub>b</sub>			11769
		TACACTAAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATT	774
Q <sub>y</sub>			
		TATATTAAATATATATTA	11770
D <sub>b</sub>			11829
		AATTTTAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTTA	834
Q <sub>y</sub>			
		TAAATATATAAAAAACATATAATTA	11830
D <sub>b</sub>			11889

RESULT 87

RES001 87  
US-09-949-016-15509  
; Sequence 15509, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CU001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15509  
; LENGTH: 134987  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15509

	Query Match	7.3%;	Score 63.6;	DB 3;	Length 134987;
	Best Local Similarity	44.5%;	Pred. No. 0.0024;		
	Matches 294;	Conservative	0;	Mismatches 364;	Indels 2; Gaps 1;
QY	177 CTACACTTTAAACAATTTAATCCAAAATCAAGATAATGACCGAACTAATTATGTGCTC	236			
Db	11230 CGACAGAGCTAGACTTCCTCAAASAAAAAAGGAATCTCTTGGTTTTATAT	11289			
QY	237 TAATTTACAACCAAGCTCATCTAGAAAATGAAAATTTTGTA AAAACCTTTTACGATTATTT	296			
Db	11290 ATATTTTTTTTATATATAATATATATATAATATAATATATATATATAATAA	11349			
QY	297 TCCTGANGCTCATTTGGGATATGATTTTTTTTAAACAACCTTTAAAGAATTTTAATGCCTATTT	356			
Db	11350 TATATAAATATATATATATATATATATTTTTTATATATATATATATATATATATATATATAT	11409			



US-09-949-016-17600

```
Query Match          7.2%; Score 63.2; DB 3; Length 50383;
Best Local Similarity 44.9%; Pred. No. 0.0026;
Matches 323; Conservative 0; Mismatches 393; Indels 4; Gaps 2;

Qy 156 TAAATTTCTCTTTGAGCAATACTACATTTTAAACATTTAAATCCAAATCAAGAAATATGA 215
    |||||
Db 33058 TAAATTTACATTAATATGCTTATATAAATATATATATATATATATATATATATATATAT 33117

Qy 216 GACCGAATTAATATGTTCTTAATTAACAACAAGCTCATCTAGAGAAATGAAATTTTGT 275
    |||||
Db 33118 AATCTATTTAAATATATAT--AAATATCATCTATTTAAATATATAAATATATATAAAT 33175

Qy 276 AAAAACTTTTACGATATATTTCTCTGATGCTCAITTTGGGATATGATTTTTTAAACAACT 335
    |||||
Db 33176 AGATGATTTTAAATATATAAATTTATATATAAATAGATGATTTTAAATATATATAAAT 33235

Qy 336 TAAAGAAATTTAATGCTTATTTTAAATTTTCAAGAAATTTTCAATCAAGAAATTTACCTC 395
    |||||
Db 33236 TATATATTAATAAATGATTTTAAATTTATATATATATATATATATATATATATATATAT 33295

Qy 396 AGGGGTCTATATGTCAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTCCGG 455
    |||||
Db 33296 ATATATAAATATATGTTTAAATAGATGTTTAAATATATATAAATTTATGTTT-- 33353

Qy 456 AATTGATTTTATCAAAATGGGTCACTTTATGCTTTTGCATACCAACAAGAAATCTTTT 515
    |||||
Db 33354 AAATAGATGTTTAAATATATATAAATTTATGTTTAAATAGATGTTTAAATATATA 33413

Qy 516 AAAAAGTCCCTGATTTTAAATATGATCGCTCGCACATATATCGACATATGTAATAAATAC 575
    |||||
Db 33414 TAAATGATGTTTAAATAGATGTTTAAATATATATAAATTTATATATATAAATTTAT 33473

Qy 576 AGATATAAAGCTTTAGAAATTTCTAGAAAAAATTTTACAAAAAATGTTTATGCTTATG 635
    |||||
Db 33474 ATTTAAATAGATCTATTTAAATATATATAAATTTATATATAAATTTATATATATAAATAG 33533

Qy 636 TCCATAATGCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTAAATTTCAAAATTTAT 695
    |||||
Db 33534 ATCTATTTAAATATATATAAATTTATATATATATATATATATATATATATATAAATTT 33593

Qy 696 CATACAGAAAAAATAACTACATCAAGATATATCTACTACTCTAGTAGGGCTTATGG 755
    |||||
Db 33594 ATATATTTAAATAGATCTATTTAAAGTATATATAAATTTATATATAAATAGATCTATTTA 33653

Qy 756 AAAATTTTCAAAAAATATTAATTTTAAAAAAATTAAGAAATGTTTATTAACA 815
    |||||
Db 33654 AATATTTATAAATTTATGTTTAAATAAATCTATTTAAATTTTATAAATTTATGTTATTA 33713

Qy 816 GTTCATAAAGATCTATTAAGATTAACCTTAGTGATATAAAGCATTTATTTCAAGGAAATA 875
    |||||
Db 33714 AATAAATCTATTTAAATATTTTATAAATTTATGTTTAAATAAATCTATTTTAAATATATA 33773
```

RESULT 90

```
US-09-949-016-17600/c
; Sequence 17600, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIORITY FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIORITY FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 17600
; LENGTH: 50383
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17600
```

```
Query Match          7.2%; Score 63.2; DB 3; Length 50383;
Best Local Similarity 44.0%; Pred. No. 0.0026;
Matches 362; Conservative 0; Mismatches 453; Indels 7; Gaps 2;
```

```
Qy 61 CTACCAATGATTTTGTATGATATTTAGATGTAATCAATTTTATTTTGAAGATATAACTAT 120
    |||||
Db 33800 CAAAAAGGGAATAGTTCTTATTTAAATATATATATATAAATAGATTTATTTAAATACATA 33741

Qy 121 CTTGTGTAATAAATGCAAGACGCTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
    |||||
Db 33740 ATTTTATTAATATTTTAAATAGATTTTATTTAAATACATAAATTTTAAATATTTAAATAGATT 33681

Qy 181 ACTTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240
    |||||
Db 33680 TATTTAAATACATAATTTTAAATATATATAAATTTTAAATAGATCTATTTAAATATATAAATTTATATA 33621

Qy 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTTCGAAATTTTTCCT 300
    |||||
Db 33620 TACTTAAATAGATCTATTTTAAATATATATAAATTTTATATAAATTTTAAATAGATCTATTTAAAT 33561

Qy 301 GA-----TGCTCATTTGGGATATGATTTTTTAAACAACCTTTAAAGAAATTTAATGCTTAT 354
    |||||
Db 33560 ATATATTTTATATATATTTTAAATAGATCTATTTTAAATATATAAATTTTATATAAATTTTAAATTTA 33501

Qy 355 TTTTAAATTTTCAAGAAATTTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTGTGCA 414
    |||||
Db 33500 TATATTTTAAATAGATCTATTTTAAATATATAAATTTTATATAAATTTTATATAAATTTTAAATTTTAA 33441

Qy 415 GTAGCCATAGCCCTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAAT 474
    |||||
Db 33440 AATACATCTAATTAATACATCATTTATATATATATATATATTTTAAATACATCTATTTTAAATACATAA 33381

Qy 475 GGGTCATCTTATGCTTTTGTATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTT 534
    |||||
Db 33380 TTTTATATATATTTTAAATACATCTATTTTAAATACATAAATTTTATATATATTTTAAATACATC 33321

Qy 535 AAAAATGATCGCTCGCACATATATCGACATAGTAAATATACAGATATAAAGCTTTTAGAA 594
    |||||
Db 33320 TATTTTAAATACATAAATTTTATATATATTTTAAATACATCTATTTTAAATACATTTATATAAAT 33261

Qy 595 TTTCTAGAAAAAATTTTACAAAAATAAACTATAT-TGCTTATGTCCTAAATAGTCTTTTAGC 653
    |||||
Db 33260 TTAATACATCTTATTTTAAATATATAAATTTATATATATATTTTAAATACATCTATTTTAAATATAT 33201

Qy 654 AAAATTTTATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATCAAGAAAAATAA 713
    |||||
Db 33200 ATAATTTTATATTTTAAATACATCTATTTTAAATATATAAATTTTATATAAATTTTAAATAGATGATA 33141

Qy 714 CTACACTAAGATATATCTACTACTCTAGTAGGGCTTATGGAATTTTTCAAAAATAT 773
    |||||
Db 33140 ATTTTATATATTTTAAATAGATTTTATTTTAAATAGATTTTATTTTAAATATATAAATTTATATATTTT 33081

Qy 774 TAAATTTTAAAAAATAAAAAATTTTAAAGAAATGTTTATTTTACAACTTGTATAAAGATCTATT 833
    |||||
Db 33080 ATAAGCAATATTAATGTAAATTTTACTATGAATCATTTATTTATTTATTTATTTATTTATTTAATA 33021

Qy 834 AAGATTACCTAGTGATATAAAGCATTTATTTTCAAGGAAAAATA 875
    |||||
Db 33020 TACAATAACATATATAAATTTTGTAAATACTATAAATAACATA 32979
```

RESULT 91

```
US-09-902-540-1280
; Sequence 1280, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
```

```
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIORITY FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIORITY FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1280
; LENGTH: 1039
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1039)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1280

Query Match          7.1%; Score 62; DB 3; Length 1039;
Best Local Similarity 47.5%; Pred. No. 0.003;
Matches 182; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

Qy 494 ATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAATGATCGCTCGCACT 553
Db 198 AAAAAAAAAAAAAAAAAACATTTATAAAAAAAAAAAAAAAAAATATATAAAAAACAT 257

Qy 554 ATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTACA 613
Db 258 ACATAAAAAAAAAAAAAAAAAATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 317

Qy 614 AATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAATTTTATAGAACTAGCGC 673
Db 318 AAAAAAAAAATTAATCAATCATCAATTAATAAATAAATAAATAAATAAATAAATAAATAA 377

Qy 674 CAATTTAAATTTCAATTTTATCATACAAGAAAAAATACTACACTTAAAGATATATCA 733
Db 378 AATTAATAAATAAATAAATAAATTAATTTACAAACAAATAAAGAAAAATAAATAAATAAATA 437

Qy 734 TACCTTCTAGTGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTCAAAAAATAAAAA 793
Db 438 TAATGTGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 497

Qy 794 TTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTACCTAGTGATATA 853
Db 498 AATGAAAAACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 557

Qy 854 AGCATTTTCAAGGAAATATA 876
Db 558 TAAAAAAGATATAAATAAATAAATA 580

RESULT 92
US-09-949-016-14164
; Sequence 14164, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIORITY FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIORITY FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 14164
; LENGTH: 18773
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14164

Query Match          7.1%; Score 62; DB 3; Length 18773;
Best Local Similarity 43.6%; Pred. No. 0.004;
Matches 343; Conservative 0; Mismatches 430; Indels 13; Gaps 1;

Qy 90 TAATCAATTTTATTTTGAAGATAAATACTATCTCTGGTAAAAAATCAAGACGATGTTTATA 149
Db 17231 TTATATATAATTAATAAATAAATAATAAGATAATCTAATAATGTAATAATAATAACATA 17290

Qy 150 CACCCCTAATTTCTTTTGAGCAATACTACACTTTTAAACATTTAAATCCAAAAACAAGA 209
Db 17291 TAAAAATAATTTTAAATATAATATTTTATATAATAATTTTATATAATAATAATAATAAT 17350

Qy 210 ATATGAGACCGAACTAATTTATGTGTTCTTAATTAACAACCAAGCTCATCTAGAAAAATGAAA 269
Db 17351 ATATATTTTATATAATAATTTAATAATTAATAATAATAATAATAATAATAATAATAATA 17410

Qy 270 TTTTGTAATAAATTTTACGATTAATTTTCCCTGATGCTCATTTGGGATATGATTTTTTAA 329
Db 17411 ATTATTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17470

Qy 330 ACACTTTAAAGAAATTTAATGCTTATTTTAAATTTTCAACGAAATTTTATTTCAATCAAGAAT 389
Db 17471 AATTAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17530

Qy 390 TACCTCAGGGTCTATATGTGTCAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCT 449
Db 17531 ATATAATAATAATAATGTTATCTATATATATATAATAATAATAATAATAATAATAATA 17590

Qy 450 TTCGGGAATGATTTTATCAAAATGGGTCATCTTATGCTTTTGATACCAAAACAAGAAA 509
Db 17591 TATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17650

Qy 510 TCTTTTAAACTAGCCCTGATTTTAAATAATGATCGCTCGCACTATATCGGACATAGTAA 569
Db 17651 TAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17697

Qy 570 AATACAGATATAAAGCTTTTAGAATTTCTAGAAAAAACTTACAAAAATAAACTATATTG 629
Db 17698 AATATAATGTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17757

Qy 630 CTTATGCTCAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCAAAATTTAAATTTCAA 689
Db 17758 TATATATATTATATGTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17817

Qy 690 TTTTATCATACAAGAAAAATACTACATAAAGATATATCTACATACCTTCTAGTGAGGC 749
Db 17818 TATATAATTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17877

Qy 750 TTATGGAATAATTTTCAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATA 809
Db 17878 TAATATAATTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17937

Qy 810 TTACAAGTTGATAAAGATCTTATTAAGATTACCTAGTGATATAAAGCAATTTTCAAAGG 869
Db 17938 TTTTATATAATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 17997

Qy 870 AAAATA 875
Db 17998 AATAATA 18003

RESULT 93
US-09-949-016-12699/c
; Sequence 12699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```









; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37165  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-37165

Query Match 6.9%; Score 60.8; DB 3; Length 601;

Best Local Similarity 49.8%; Pred. No. 0.0049;  
Matches 207; Conservative 0; Mismatches 207; Indels 2; Gaps 2;

```
QY 462 TTTTATCAAAATGGTCTATCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAACT 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 593 TTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 522 AGCCCTGATTTTAAATATGCTCGCCTATATCGGACATAGTAAAAATACAGATAT 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 533 ATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 582 -AAAAGCTTTAGAAATTTCTAGAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTA 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 473 AAAAATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ATAGCTTTTAGCAAAATTT-TATAGAACTAGGCCCAAAATTTAAATTTCAAATTTTATCATA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 473 AAAAATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ATAGCTTTTAGCAAAATTT-TATAGAACTAGGCCCAAAATTTAAATTTCAAATTTTATCATA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 473 AAAAATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 820 ATAAAGATCTATTAAAGATTACCTAGTGATATAAGCATTTTCAAAGGAAATA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 233 AAATATATAAAATATATTTATATATTTTAAATATAAATATAAATATAAATATAAATAT 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 100

US-09-949-016-145869/c  
; Sequence 145869, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 145869

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-145869

Query Match 6.9%; Score 60.8; DB 3; Length 601;

Best Local Similarity 49.8%; Pred. No. 0.0049;  
Matches 207; Conservative 0; Mismatches 207; Indels 2; Gaps 2;

```
QY 462 TTTTATCAAAATGGTCTATCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAACT 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 593 TTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 522 AGCCCTGATTTTAAAAATGATCGCTCGCCTATATCGGACATAGTAAAAATACAGATAT 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 533 ATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 582 -AAAAGCTTTAGAAATTTCTAGAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTA 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 473 AAAAATATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ATAGCTTTTAGCAAAATTT-TATAGAACTAGGCCCAAAATTTAAATTTCAAATTTTATCATA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 413 ATAAAAATATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAAT 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 CAAGAAAAAATAAATCTACACTAAAGATATACCTCATACCTTCTAGTGAGGCTTATGGAATA 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 353 ATATAAATAATATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAA 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 760 TTTTCAAAAAATATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 293 ATATATAAATAATATATATTTTAAATATAAATATAAATAAATAAATAAATAAATAAATAA 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 820 ATAAAGATCTATTAAAGATTACCTAGTGATATAAGCATTTTCAAAGGAAATA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 233 AAATATATAAAATATATTTATATATTTTAAATATAAATATAAATATAAATATAAATAT 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 101

US-09-949-016-146137/c

; Sequence 146137, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 146137

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-146137

Query Match

Best Local Similarity 6.9%; Score 60.8; DB 3; Length 601;

Matches 207; Conservative 0; Mismatches 207; Indels 2; Gaps 2;

```
QY 462 TTTTATCAAAATGGTCTATCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAACT 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 593 TTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 522 AGCCCTGATTTTAAAAATGATCGCTCGCCTATATCGGACATAGTAAAAATACAGATAT 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 533 ATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAT 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 582 -AAAAGCTTTAGAAATTTCTAGAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTA 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 473 AAAAATATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATAT 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 ATAGCTTTTAGCAAAATTT-TATAGAACTAGGCCCAAAATTTAAATTTCAAATTTTATCATA 699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 413 ATAAAAATATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAAT 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 CAAGAAAAAATAAATCTACACTAAAGATATACCTCATACCTTCTAGTGAGGCTTATGGAATA 759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





Qy 530 ATTTTAAATGATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTT 589  
|||||  
Db 82510 ATTTTCCAAATTAAAGTTTCATAATGTTTCATTTCTTTTATTTTCTTAATTAATGTT 82569  
|||||  
Qy 590 TAGAATTTCTAGAAAAAATTTACAAAAATAAACTATATGCTTATGCTCCTAATAGTCTTT 649  
|||||  
Db 82570 TTTATATATTAAATAAAATTAATTAATAAAGTAAGCACCATAAGATATATTGTGCAG 82629  
|||||  
Qy 650 TAGCAATTTTATAGAACTAGCGCAATTTAAATTTCAAAATTTTATCATACAAGAAAAA 709  
|||||  
Db 82630 TAGTAAGCTTTATTTTACCTTAGTTAACTTTCTATCGTTTAAAGATACCAGTGAATC 82689  
|||||  
Qy 710 ATAACTACACTAAAGATATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAA 769  
|||||  
Db 82690 ATTTCCACCAAGCAAGCATGTTCTTTGTAATGCAACATTAATCTAATTAATAA 82749  
|||||  
Qy 770 ATATTAAATTTTAAAAAATAAAAAATTAAGAAAAATGTTTATTAAGTTGATAAAGATC 829  
|||||  
Db 82750 ATAAATTTATGAACAATAAATTAACAGTAAATTTATTTGTTGGTCCCAAGTAGATGAAAAA 82809  
|||||  
Qy 830 TATTAAGATTACCTAGTCATATATAAGCATTTATTTCAAGGAA 871  
|||||  
Db 82810 TTAATGTTTAGAAAGTTATATAAAGAAAGCATTTAAAAA 82851  
|||||

## RESULT 107

US-08-487-826B-13  
; Sequence 13, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chinnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellem, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobs Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH121.001CP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19124 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 6.9%; Score 60.2; DB 2; Length 19124;  
Best Local Similarity 46.8%; Pred. No. 0.009;  
Matches 257; Conservative 0; Mismatches 288; Indels 4; Gaps 2;  
Qy 320 ATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAATTTTCAGGAATTTATTTC 379  
|||||  
Db 15360 ATATATGAGTACTTGTATGTTTATATATATATATATATATATATATATATATAT 15419  
|||||  
Qy 380 ATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGCATAGCCCTAGGATACAAAG 439  
|||||  
Db 15420 AAAAGAAATATATAAAACAAATTTTAAATGAAAAAAGAAAAATGAATATAAAAA 15479  
|||||  
Qy 440 AAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAATGGGTCTTATGCTTTTGATACCA 499  
|||||  
Db 15480 AAATTTTATAAATAAAAAAAGAAAAAAGG--AGAAAAATTTTTTAAAAAAT 15536  
|||||  
Qy 500 AACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAATGATCGCTCGCACTATATCG 559  
|||||  
Db 15537 AATAAAAAATTAATAAAAAATATAAAATTTTGATAGAAATAAAAAATGAAAAAGATTATCAA 15596  
|||||  
Qy 560 GACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTTACAAAAATA 619  
|||||  
Db 15597 AAAAAATTAATAAAAAAATTTTATATAAAAAAATGATTATAAAAAAATAAAAACAA 15656  
|||||  
Qy 620 AACTATATGCTTATGTCCTCAATAGTCTTTTAGCAATTTTAT-AGAACTAGCCCAAT 678  
|||||  
Db 15657 AGAAGAAAAAACAATTAATAAAAAAATATATATATATATATATATATATATATATAT 15716  
|||||  
Qy 679 TTAATTTCAATTTTATCATACAAGAAAAAATACTACCTAAAGATATATCTCATACCT 738  
|||||  
Db 15717 AAAAAAATATATTAATAAATAATATATATATATATATATATATATATATATATATAT 15776  
|||||  
Qy 739 TCTAGTGAGCTTATGAAAAATTTTCAAAAAATTAATTTTAAAAAATAAAAAATTTAA 798  
|||||  
Db 15777 AAAAAAATAATATATACATAAAAAAATTTTATTTTAAATAAAAAAATAATAAT 15836  
|||||  
Qy 799 GAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTACCTAGTATATAAAGCAT 858  
|||||  
Db 15837 AAAAAATTTAATTAATAAAAAAATAATAATAAAAAATTAATAATAAAAAA 15896  
|||||

## RESULT 108

US-08-213-419B-3  
; Sequence 3, Application US/08213419B  
; Patent No. 633406  
; GENERAL INFORMATION:  
; APPLICANT: Inselburg, J. et al.  
; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM  
; FILE REFERENCE: JII-002CNCPC  
; CURRENT APPLICATION NUMBER: US/08/213,419B  
; CURRENT FILING DATE: 1994-03-14  
; PRIOR APPLICATION NUMBER: US 07/870,506  
; PRIOR FILING DATE: 1992-04-17  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 6124  
; TYPE: DNA  
; ORGANISM: Plasmodium falciparum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2407)..(2439)  
; NAME/KEY: CDS  
; LOCATION: (2598)..(3404)  
; NAME/KEY: CDS  
; LOCATION: (3580)..(3720)  
; NAME/KEY: CDS  
; LOCATION: (3850)..(5835)

US-08-213-419B-3

Query Match 6.8%; Score 59.8; DB 3; Length 6124;  
Best Local Similarity 52.2%; Pred. No. 0.0096;  
Matches 133; Conservative 0; Mismatches 122; Indels 0; Gaps 0;  
  
Qy 560 GACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTTACAAAAATAA 619  
Db 2124 GAGATGTACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2183  
  
Qy 620 AACTATATGCTATGCTCTTAATAGTCTTTTATGCAAAATTTTATGAACTAGCGCCAAAT 679  
Db 2184 AATAATATAATATATTTTAAATTTTATATATATATATATATATATATATATATATAT 2243  
  
Qy 680 TAAATCCAAATTTTATCATACAGAAAAATAAATAAATAAATAAATAAATAAATAAATAA 739  
Db 2244 TAAATTTAAATTTAGATTTGTCACAAAAAATAAATAAATAAATAAATAAATAAATAA 2303  
  
Qy 740 CTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTTTAAAAAATAAATAAATAAATAA 799  
Db 2304 AAAATACATATATTTAAACATAAAGAAAAATTTAAATAAATCAAAACATATTTCAAAAAAT 2363  
  
Qy 800 AAAATGTTTATTTACA 814  
Db 2364 AAAGTCTTAAAAATA 2378

RESULT 109  
US-09-662-254B-25  
; Sequence 25, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1XCI  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-662-254B-25

Query Match 6.8%; Score 59.8; DB 3; Length 50000;  
Best Local Similarity 47.0%; Pred. No. 0.012;  
Matches 395; Conservative 0; Mismatches 427; Indels 19; Gaps 6;  
  
Qy 41 AAGAAATTTGATTTATCAAGGCTACCAATGATTTTGTATTTAGATGTTAATCAATTTT 100  
Db 10060 AATATTTGTAATTTTATAAACAATAAATAAATAAATAAATAAATAAATAAATAAAT 10119  
  
Qy 101 ATTTTGAAGATAAATACTATCTTGGTAAAAAATGCAAGAGTGTTTTA-----CACCCCT 156  
Db 10120 TATTATATCTGTATTTGTAACAACTTAAATATTTAAATTTATTTTTCATATATACATC 10179  
  
Qy 157 ATTTCTCTTTGAGCATATCTACACTTTTAAACATTTTAACTTAACTTAACTTAACTTAA 216  
Db 10180 AATATTTCTTACACACATCTTCCATTAACATATTTCTTCTTCTATATATCGCATTTTAT 10239  
  
Qy 217 ACCGAACATTAATGATGTTCTTAAATACAAACCAAGCTCATCTAGAAAAATGAAAAATTTGTA 276  
Db 10240 ATCTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 10299  
  
Qy 277 AAAACTTTTACGATATTTTCTCTGATGCTCATTTGGGATATGATTTTTTTTAAACAACTT 336

Db 10300 AATTATCAATGCTTTTAAATTTTCATGATAATTT-----ACACATATATAATATATA 10352  
  
Qy 337 AAAGAATTTAATGCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 396  
Db 10353 TATATATATAATATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 10411  
  
Qy 397 GGGGCTATATGTTGTCAGTAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAGCCATAG 456  
Db 10412 GTAATCGATCTATTTTGAATAATATAATTAATTAATTAATTAATTAATTAATTAATTAAT 10471  
  
Qy 457 ATTGATTTTATCAAAATGGGTCACTTATGCTTTTGTATGATCAAAATCAAAATCAAAATCTTT 516  
Db 10472 TTTTAAATGGAATTTATGCG--AAATATTTATTTATATAAAGAAAAAATATATTTATGA 10529  
  
Qy 517 AAATAGCCCTGATTTTAAATAATGATCGCTCGCACTATATCGGACATAGTAAAAATACA 576  
Db 10530 TATGTTAGAAATATTTTATTTAAATTTTGTGTTTACATTTACTATATAGTTTTTAAAAA 10589  
  
Qy 577 GATATAAAGCTTTTGAATTTCTAGAAAAAATTTTCAAAATAAATCTATATTTCTTATGT 636  
Db 10590 TATAAAATATTTATGATTAATGATTTGATTAATTAATTAATTAATTAATTAATTAATTA 10649  
  
Qy 637 CCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCA-AAFTTAAATTTCAAATTTTAT 695  
Db 10650 TATA-----TATTATCCAAAAAGTAAATATCTATAAAATTTATAGAAATTTAGCCAA 10705  
  
Qy 696 CATACAAGAAAAATACTACACTAAAGATATATCTATCTTCTAGTGAGGCTTATGG 755  
Db 10706 AATTAAAAATAAATAATTTATGTTTCAAGCAACGAAAAATGAATAATTTAAATTTTATG 10765  
  
Qy 756 AAAATTTTCAAAATAATTAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAA 815  
Db 10766 ATACTAATGAACAATAATTTATTTTAAATAATTAATTAATTAATTAATTAATTAATTAAT 10825  
  
Qy 816 GTTGATAAAGATCTTATTAAGATTTACTAGTGATATAAAGCAATTTTCAAAGGAAAAATA 875  
Db 10826 TAAGATTTAATAATTAATGAAATAACGATAGATTTAATTAATTAATTAATTAATTA 10885  
  
Qy 876 A 876  
Db 10886 A 10886

RESULT 110  
US-09-662-254B-23/c  
; Sequence 23, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1XCI  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 23  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-662-254B-23

Query Match 6.8%; Score 59.4; DB 3; Length 50000;  
Best Local Similarity 46.2%; Pred. No. 0.014;  
Matches 381; Conservative 0; Mismatches 431; Indels 13; Gaps 5;



```
;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) _ (201529)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-12740

Query Match      6.8%; Score 59.4; DB 3; Length 201529;
Best Local Similarity 47.8%; Pred. No. 0.016;
Matches 203; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

QY 442 ATTATCTTTCCGGGAATGATTTTATCAAAATGGGTCACTCTTTATGCTTTTGATACCAA 501
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61846 ATATRAAATTTATATAAATTTATATTTATATAAATTTATGTAATTTATATAATATACAATA 61787
QY 502 CAAGAAAATCTTTTAAACTAGCCCTGATTTTAAAAATGATCGCTGCACATATATCGGA 561
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61786 TATTATAAATTTTAAATTTAAATATAAATCTTAAATATAAATTTAAATATAAATTA 61727
QY 562 CATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAAA 621
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61726 ATTATATTTCTATTAATTTACATTTATATAAATTTATATTTATATTTATATTTATAAAT 61667
QY 622 CTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTTATGAAGAACTAGCGCCAAATTTA 691
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61666 ATAAATTTATATTTATATTTTATAAATATAATTTATATTTTATATAATATAAATTTATA 61607
QY 682 AATTCAAATTTTATCATACAAGAAAAATAACTACACTAAGATATATCTCATACCTTCT 741
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61606 TTTATAAATTTATATAATATAAATATAATTTTATATTTTATATTTTATATAAATAAAT 61547
QY 742 AGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTTAAAAAATAAAAAATTTAAAGAA 801
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61546 ATATTTATATTTAAATTTAAATTTAAATATAAATATAATTTATATAATATAAATAAATA 61488
QY 802 AATGTTTATTAACAAGTTGATAAAGATCTATTAAGATTTACCTAGTGATATAAAGCATTTAT 861
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61487 TATATAAATAAATAATATAATTAATATAATAGTATAAATATAGTTTTTTTATATATAT 61428
QY 862 TTCAA 866
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61427 AACTA 61423

RESULT 113
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 593827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; ADDRESS: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESS: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
```

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOHETICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match      6.7%; Score 59; DB 2; Length 19124;
Best Local Similarity 46.1%; Pred. No. 0.015;
Matches 357; Conservative 0; Mismatches 400; Indels 18; Gaps 4;

QY 68 ATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACATCTCTGGTA 127
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1553 ATGTTCTACGTTTCTCTCTAATTTTATCTAGGTGCTTACTAATAATTAATTTTATGTC 1494
QY 128 AAAAAATCAAGCAGTCTTTTACACCCCTAATTTCTTTTGGAGCAATACACTTAA 187
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1493 AAAAAATGTTTGAAGTTTTTTTAAAGATTTTATTTTATTTTAAATCTAAATCATATAAT 1434
QY 188 AACATTTAAATCCAAAAATCAAGAATATGAGCCGAATTAATTTATGTTCTTAATTCAACC 247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1433 TATGAAAAAATAATATTAATTTT-----TATTATTATTTAAAAAATAAAAAA 1386
QY 248 AAGCTCATCTAGAAAAATGAAAAATTTGTAAAAAATTTTACGATTTTCTCTGATGCTC 307
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1385 AGAGAGAACAAATTAATAAATAAAGATATATATATATATATATCTATTAGATACCAGATTGA 1326
QY 308 ATTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAAGCTTATTTTAAATTTCAAG 367
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1325 TAGTTTCTTTAGTCTTGATGTTTATACAAATATATTTTATTAATAATTTTACATATTATA 1266
QY 368 AATTTTATTTCAATCAAAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCCATAGCCC 427
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1265 ATATTTTGATAACATATCTAATATAT-ATATATATATATATATATATATATATATTTTAA 1207
QY 428 TAGGATACAAAGAAATTTATCTTTTCGGGAATGATTTTATCAAAATGGGTCATCTTAG 487
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1206 TAATAGCTATAGAATTTTATCTTATATAATTTGCAATTTATACATAACATAACAAAAATA-- 1149
QY 488 CTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATGATCGCT 547
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1148 -ATTAATTACTATATATATATTTTAAATATTTTATATATATATTTTGAAGCGTTATAC 1090
QY 548 CGCACTATATCGGCACATAGTAAAAATACAGATATAAAGCTTTA--GAATTTCTAGAAAA 605
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1089 AACATTTATATATCTATGATAATACAGCGGATAATACATTTTATTTTATATATATATTA 1030
QY 606 AACTTACAAAAATAAACTATATTTCTTATGCTCTTAATAGTCTTTTGAAGCAATTTTATAGA 665
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1029 TAATTTATATATATAAGTAATATATTTTAAATGTTGTATACATTATGACAACTTAAATGT 970
QY 666 ACTAGCCCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAAATACACTACATAAGA 725
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
969 TGTGTGAAAAATGTAATGATGAAAAATTTGTTTATTTATATATATTTAAAAATAAAAAATTTT 910
QY 726 TATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATAATTTTAAAAAATA 785
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
909 AATAAATTTTAAATGTCGATGATATGCTATAAAAATGAATAAATGAAGAAATATTTTAAAG 850
QY 786 AATAAAAAATTAAGAAAAATGTTTATTTACAAAGTTGATAAAAAGATCTTATTAAGATTA 840
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
849 AATCTAATTAATTTGATTTGAAAAATAAACCATCGAATATATTTTGAATATATAA 795
```





Db 15547 TATATTTTAATATTAG-TAAATAAATTTAATAATATGAACGTAAATAATTTAGAAAAATA 15489  
Qy 687 AAATTTTATCATACAGAAAAAATAACTACATAAGATATACTCTACTCTTAGTGA 746  
Db 15488 TCAACATTTTAAAGAAGATAATATATCATATTTAATAATATATTTATATATATACATAGA 15429  
Qy 747 GGCTTATGGAATTTTCAAAAAATATTAAATTTTAAAAAATAAAAAATTTAAAGAAAAATGT 806  
Db 15428 AGAATATATTAT 15369  
Qy 807 TTATTACAAGTTGATAAAAGATCTATTAAAGA 837  
Db 15368 ACAAATATATGTTAT 15338

RESULT 116  
US-09-949-016-12386  
; Sequence 12386, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12386  
; LENGTH: 263693  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12386

Query Match 6.6%; Score 58.2; DB 3; Length 263693;  
Best Local Similarity 51.1%; Pred. No. 0.028;  
Matches 164; Conservative 0; Mismatches 153; Indels 4; Gaps 1;  
Qy 555 TATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTAGAAAAAACTTACAA 614  
Db 174466 TATATAATATAATAATATATATCTTTTATATATATATATATATATATATATATATAT 174525  
Qy 615 AATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCC 674  
Db 174526 CTAT 174585  
Qy 675 AAATTTAAATTCAAATTTTATCACAAGAAAAAATACTACATAAGATATACCTCAT 734  
Db 174646 AAATTTAT 174645  
Qy 735 ACCTTCTAGGAGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTCAAAAAATTTTCA 794  
Db 174646 AAATTTAT 174705  
Qy 795 TAAAGAAAATGTTTAA-----TTACAAGTTGATAAAGATCTATTAAGATTACCTAGTGATA 850  
Db 174706 TATAGTTTATATATATATGTTTAAATATATATATATATATATATATATATATATATAT 174765  
Qy 851 TAAAGCATTTATTTCAAGGAA 871  
Db 174766 TAAAGTATGAAGAAAAATGAA 174786

RESULT 117  
US-09-949-016-16915  
; Sequence 16915, Application US/09949016

; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16915  
; LENGTH: 263694  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16915

Query Match 6.6%; Score 58.2; DB 3; Length 263694;  
Best Local Similarity 51.1%; Pred. No. 0.028;  
Matches 164; Conservative 0; Mismatches 153; Indels 4; Gaps 1;  
Qy 555 TATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTAGAAAAAACTTACAA 614  
Db 174466 TATATAATATAATAATATATATCTTTTATATATATATATATATATATATATATATAT 174525  
Qy 615 AATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCC 674  
Db 174526 CTAT 174585  
Qy 675 AAATTTAAATTCAAATTTTATCACAAGAAAAAATACTACATAAGATATACCTCAT 734  
Db 174586 AATAAAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 174645  
Qy 735 ACCTTCTAGGAGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTCAAAAAATTTTCA 794  
Db 174646 AAATTTAT 174705  
Qy 795 TAAAGAAAATGTTTAA-----TTACAAGTTGATAAAGATCTATTAAGATTACCTAGTGATA 850  
Db 174706 TATAGTTTATATATATATGTTTAAATATATATATATATATATATATATATATATATAT 174765  
Qy 851 TAAAGCATTTATTTCAAGGAA 871  
Db 174766 TAAAGTATGAAGAAAAATGAA 174786

RESULT 118  
US-09-949-016-15014/c  
; Sequence 15014, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15014  
; LENGTH: 55264  
; TYPE: DNA



;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
;; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
;; TITLE OF INVENTION: AND USES THEREOF  
;; FILE REFERENCE: CL000790  
;; CURRENT APPLICATION NUMBER: US/09/949,002  
;; CURRENT FILING DATE: 2000-01-28  
;; PRIOR APPLICATION NUMBER: 60/231,401  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 10823  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 627  
;; LENGTH: 105001  
;; TYPE: DNA  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (1)...(105001)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-949-002-627

Query Match 6.6%; Score 57.6; DB 3; Length 105001;  
Best Local Similarity 33.0%; Pred. No. 0.034;  
Matches 237; Conservative 0; Mismatches 482; Indels 0; Gaps 0;  
QY 157 AATTTCCTCTTTGAGCAATACACACTTTAAACACATTTAATCCAAATCAAGAATATGAG 216  
DB 32439 AACGTATATATTTGTAAATAACATATAAACGTATATATTTGTAAATAACATATAAA 32380  
QY 217 ACCGAACTAATATGTGTCTTAATACCAACGCTCATCTAGAAATGAAAAATTTTGTGA 276  
DB 32379 CGTATATATTTGTAAATAACATATAAACGTATATATTTGTAAATAATATAAACGTA 32320  
QY 277 AAAACCTTTTACGATATTTTCTGATGCTCATTTGGGATATGATTTTAAACAACATT 336  
DB 32319 TATATTTGTAAAAATATAAACGTATATATTTTGTAAATATAATAACGTATATATTT 32260  
QY 337 AAAGAATTTAATGCTTATTTTAAATTTCAAGAAATTTTCAATCAAGAATATACCTCA 396  
DB 32259 TTGTAATAATATAAAGTATATATATTTGTAAATATAATAAACGTATATATTTGTA 32200  
QY 397 GGGGTCTATGTGTGAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGA 456  
DB 32199 AATATATAAATATATTTGTAAATATAATAAACGTATATATTTGTAAATATAATAAC 32140  
QY 457 ATTGAATTTTATCAAAATGGGTCACTTATGCTTTTGATACCAACAAAGAAATCTTTTA 516  
DB 32139 ATATATTTGTAATATATATATTTGTAAATATAATAAACGTATATATTTGTAAATATA 32080  
QY 517 AAACCTAGCCCTGATTTTAAATGATCGCTCGCACTATATCGGACATAGTAAATAACA 576  
DB 32079 NNN 32020  
QY 577 GATATAAAGCTTTAGAATTTCTAGAAAACTTACAAAAATAAACTATATTTGCTTATGT 636  
DB 32019 NNN 31960  
QY 637 CCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTAAATCTCAATTTATC 696  
DB 31959 NNN 31900  
QY 697 ATACAAGAAAAATAACTACACTAAAGATATACTCATACCTTCTAGTAGGCTTATGGA 756  
DB 31899 NGTAAATATATATATTTTGTAAATATAATAAACGTATATATTTGTAAATATAATAACA 31840  
QY 757 AAATTTTCAAAAAATATTAATTTTAAAAAAATAAAATTTAAAGAAATGTTTATTACAAG 816  
DB 31839 TATAATTTTATAAATATATTTTATAAATATAATAAACGTATATATTTTATAAATATAT 31780  
QY 817 TTGTAAGAAATCTTAAAGATTAACCTAGTGATATAAGCAATATTTTCAAGGAAAAATA 875  
DB 31779 TATAATATATATTTGTAATATGTTTATAAATATAATAATAATAATAATAATAATA 31721

RESULT 122  
US-09-949-002-821/c  
; Sequence 821, Application US/09949002  
; Patent No. 6900016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000790  
; CURRENT APPLICATION NUMBER: US/09/949,002  
; CURRENT FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/231,401  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 10823  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 821  
; LENGTH: 105002  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(105002)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-002-821

Query Match 6.6%; Score 57.6; DB 3; Length 105002;  
Best Local Similarity 33.0%; Pred. No. 0.034;  
Matches 237; Conservative 0; Mismatches 482; Indels 0; Gaps 0;  
QY 157 AATTTCCTCTTTGAGCAATACACACTTTTAAACACATTTAATCCAAATCAAGAATATGAG 216  
DB 32439 AACGTATATATTTGTAAATAACATATAAACGTATATATTTGTAAATAACATATAAA 32380  
QY 217 ACCGAACTAATATGTGTCTTAATACCAACGCTCATCTAGAAATGAAAAATTTTGTGA 276  
DB 32379 CGTATATATTTGTAAATAACATATAAACGTATATATTTGTAAATAATATAAACGTA 32320  
QY 277 AAAACCTTTTACGATATTTTCTGATGCTCATTTGGGATATGATTTTAAACAACATT 336  
DB 32319 TATATTTGTAAAAATATAAACGTATATATATTTTGTAAATATAATAAACGTATATATTT 32260  
QY 337 AAAGAATTTAATGCTTATTTTAAATTTCAAGAAATTTTCAATCAAGAATATACCTCA 396  
DB 32259 TTGTAATAATATAAAGTATATATATTTGTAAATATAATAAACGTATATATTTGTA 32200  
QY 397 GGGGTCTATGTGTGAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGA 456  
DB 32199 AATATATAAATATATTTGTAAATATAATAAACGTATATATTTGTAAATATAATAAC 32140  
QY 457 ATTGAATTTTATCAAAATGGGTCACTTATGCTTTTGATACCAACAAAGAAATCTTTTA 516  
DB 32139 ATATATTTGTAATATATATATTTGTAAATATAATAAACGTATATATTTGTAAATATA 32080  
QY 517 AAACCTAGCCCTGATTTTAAATGATCGCTCGCACTATATCGGACATAGTAAATAACA 576  
DB 32079 NNN 32020  
QY 577 GATATAAAGCTTTAGAATTTCTAGAAAACTTACAAAAATAAACTATATTTGCTTATGT 636  
DB 32019 NNN 31960  
QY 637 CCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTAAATCTCAATTTATC 696  
DB 31959 NNN 31900  
QY 697 ATACAAGAAAAATAACTACACTAAAGATATACTCATACCTTCTAGTAGGCTTATGGA 756  
DB 31899 NGTAAATATATATATTTTGTAAATATAATAAACGTATATATTTGTAAATATAATAACA 31840  
QY 757 AAATTTTCAAAAAATATTAATTTTAAAAAAATAAAATTTAAAGAAATGTTTATTACAAG 816  
DB 31839 TATAATTTTATAAATATATTTTATAAATATAATAAACGTATATATTTTATAAATATAT 31780







```
Db 455 ATATATATTTTAAATATAAATAATAAATAATAATATATATTTTAAATATAAATATAATAA 396
Qy 739 TCTAGTGAGGCTTATGAAATTTTCAAAAAATTAATTTTAAAAAATAAATAAATAAATAA 798
Db 395 ATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAA 336
Qy 799 GAAATGCTTTTACACAGTTGATAAAGATCTATTAGATTACCTAGTAGATATAAAGCAT 858
Db 335 AAATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAA 276
Qy 859 TATTTCAAAGGAAATA 875
Db 275 ATATTAAATATTGAATA 259

RESULT 130
US-09-949-016-37152/c
; Sequence 37152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37152
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37152

Query Match 6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

Qy 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAATAA 619
Db 575 ATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAAT 511
Qy 620 AACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
Db 515 ATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAAT 456
Qy 679 TTAATTCAAATTTTATCATACAGAAAAAATAAATTAATTTTAAAAAATAAATAAATAA 738
Db 455 ATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAA 396
Qy 739 TCTAGTGAGGCTTATGAAATTTTCAAAAAATTAATTTTAAAAAATAAATAAATAAATAA 798
Db 395 ATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAA 336
Qy 799 GAAATGCTTTTACACAGTTGATAAAGATCTATTAGATTACCTAGTAGATATAAAGCAT 858
Db 335 AAATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAA 276
Qy 859 TATTTCAAAGGAAATA 875
Db 275 ATATTAAATATTGAATA 259

RESULT 131
US-09-949-016-37153/c
; Sequence 37153, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37153
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37153

Query Match 6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

Qy 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAATAA 619
Db 570 ATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAAT 511
Qy 620 AACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
Db 510 ATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAA 451
Qy 679 TTAATTCAAATTTTATCATACAGAAAAAATAAATTAATTTTAAAAAATAAATAAATAAATAA 738
Db 450 ATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAA 391
Qy 739 TCTAGTGAGGCTTATGAAATTTTCAAAAAATTAATTTTAAAAAATAAATAAATAAATAAATAA 798
Db 390 ATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAAATAA 331
Qy 799 GAAATGCTTTTACACAGTTGATAAAGATCTATTAGATTACCTAGTAGATATAAAGCAT 858
Db 330 AAATATATATATTTTAAATATAAATAATAAATAATAATATATTTTAAATATAAATAATAA 271
Qy 859 TATTTCAAAGGAAATA 875
Db 270 ATATTAAATATTGAATA 254

RESULT 132
US-09-949-016-37166/c
; Sequence 37166, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37166
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37166
```

```

; ORGANISM: Human
US-09-949-016-37166

Query Match          6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

QY 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 ATATTTTAAATATAAATATAAATAATATATATATTTTAAATATAAATATAAATAAT 511

QY 620 AACTATATGCTTATGCTCTAATAGCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAA 451

QY 679 TTAATTCAAATTTTATCATACAGAAAAAATACTACACTAAAGATATACTCATACCT 738
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 ATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATA 391

QY 739 TCTAGTGAGGCTTATGGAATAAATTTTCAAAAAATATTAATTTTAAAAAATAAATAA 798
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 ATATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAA 331

QY 799 GAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTACCTAGTAGTATAAAGCAT 858
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 AAATATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATA 271

QY 859 TATTTCAAAGGAAATA 875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 ATATTAATATTTGAATA 254

RESULT 133
US-09-949-016-37167/c
; Sequence 37167, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37167
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37167

Query Match          6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

QY 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 ATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATAATATAAATAAT 511

QY 620 AACTATATGCTTATGCTCTAATAGCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 ATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATAA 516

QY 620 AACTATATGCTTATGCTCTAATAGCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 ATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATA 456

QY 679 TTAATTCAAATTTTATCATACAGAAAAAATACTACACTAAAGATATACTCATACCT 738
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 ATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAATATA 396
```

```

QY 739 TCTAGTGAGGCTTATGGAATAAATTTTCAAAAAATATTAATTTTAAAAAATAAATAA 798
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 ATATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAA 336

QY 799 GAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTACCTAGTAGTATAAAGCAT 858
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 AAATATATATATTTTAAATATAAATAATATAAATAATATATTTTAAATATAAATA 276

QY 859 TATTTCAAAGGAAATA 875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ATATTAATATTTGAATA 259

RESULT 134
US-09-949-016-145870/c
; Sequence 145870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145870
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145870

Query Match          6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

QY 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 ATATTTTAAATATAAATATAAATAATATATATTTTAAATATAAATAATATAAATAAT 511

QY 620 AACTATATGCTTATGCTCTAATAGCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAA 451

QY 679 TTAATTCAAATTTTATCATACAGAAAAAATACTACACTAAAGATATACTCATACCT 738
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 ATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAA 391

QY 739 TCTAGTGAGGCTTATGGAATAAATTTTCAAAAAATATTAATTTTAAAAAATAAATAA 798
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 ATATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATAA 331

QY 799 GAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTACCTAGTAGTATAAAGCAT 858
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 AAATATATATATTTTAAATATAAATAATATAAATAATATATATTTTAAATATAAATA 271

QY 859 TATTTCAAAGGAAATA 875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 ATATTAATATTTGAATA 254

RESULT 135
US-09-949-016-145871/c
; Sequence 145871, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```





```
QY 799 GAAATGCTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTACCTAGTGTATATAAAGCAT 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 AAATATATATATTTTAAATATAAATATAAATAAATATATATATATATATATATATATATAT 276

QY 859 TATTTCAAAGGAAATA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ATATTATATATTGGAATA 259

RESULT 138
US-09-949-016-146406/c
; Sequence 146406, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146406
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-146406

Query Match 6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

QY 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAATAA 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 ATATTTTAAATATAAATATAAATAAATATATATATTTTAAATATAAATATAAATAATAT 516

QY 620 AACTATATTGCTTATGCTCTTAATAGCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 ATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATAATAT 456

QY 679 TTAATTCAAAATTTTATCATACAAGAAAAATAACTACACTAAAGATATACCTACCT 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 ATATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATAA 396

QY 739 TCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATATATTTTAAAAAATAAATAATAA 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 ATATATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATAA 336

QY 799 GAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTACCTAGTGTATATAAAGCAT 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 AAATATATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATAAATAA 276

QY 859 TATTTCAAAGGAAATA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ATATTATATTGGAATA 259

RESULT 140
US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
QY 799 GAAATGCTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTACCTAGTGTATATAAAGCAT 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 AAATATATATATTTTAAATATAAATATAAATAAATATATATATATATATATATATATATAT 276

QY 859 TATTTCAAAGGAAATA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 ATATTATATATTGGAATA 259

RESULT 139
US-09-949-016-146407/c
; Sequence 146407, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 146406
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-146406

Query Match 6.4%; Score 56.2; DB 3; Length 601;
Best Local Similarity 52.7%; Pred. No. 0.037;
Matches 167; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

QY 561 ACATAGTAAAAATACAGATAT-AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAATAA 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 ATATTTTAAATATAAATATAAATAAATATATATTTTAAATATAAATATAAATAATAT 511

QY 620 AACTATATTGCTTATGCTCTTAATAGCTTTTAGCAAAATTT-TATAGAACTAGCGCCAAAT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 ATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATAAATAA 451

QY 679 TTAATTCAAAATTTTATCATACAAGAAAAATAACTACACTAAAGATATACCTACCT 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 ATATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATAA 391

QY 739 TCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATATATTTTAAAAAATAAATAATAA 798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 ATATATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATATAAATAA 331

QY 799 GAAATGCTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTACCTAGTGTATATAAAGCAT 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 AAATATATATATTTTAAATATAAATATAAATAAATATAAATATAAATATAAATAAATAA 271

QY 859 TATTTCAAAGGAAATA 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 ATATTATATTGGAATA 254

RESULT 139
US-09-949-016-146407/c
; Sequence 146407, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```









```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16600
; LENGTH: 298336
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(298336)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16600

Query Match      6.1%; Score 55.2; DB 3; Length 298336;
Best Local Similarity 46.4%; Pred. No. 0.11;
Matches 247; Conservative 0; Mismatches 283; Indels 2; Gaps 2;

Qy      342  ATTATATGCTTATTTTAAATTTTCACGAAATTTTCAATCAAGAAATTACCTCAGGGGT 401
Db      273693  ATTCTATACGTGTATATACTATATATTCTATACGCTATATAATACTATATATT 273752

Qy      402  CTATATGTGTGAGTACGCCATAGCCCTAGGATCAAGAAATTTATCTTTTCGGGAATTGA 461
Db      273753  CTATACGTGTATATAATAATACTATATATTCTATACGCTATATAAATACTATATATTCTA 273812

Qy      462  TTTTATCAAAATGGGTCTATTCTTTTGATACCAACAGAAATCTTTTAAACT 521
Db      273813  TACGTGTATATACTAACTATATATTCTATACGCTGTATATAATAATACTATATATTCT 273872

Qy      522  AGCCCTGATTTTAAAAATGATCGCTCGCACTATATCGGACATAGTAAAAATACAGATAT 581
Db      273873  ATACGTGTATATAATAATAACAGTATATAATAATACTATATATAATACTATA 273932

Qy      582  AAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACTATATTGCTTTATGCTCAA 641
Db      273933  TATTATATACACGTATATAACATATATTATATAATAATAATAATAATAATAATAATAATA 273992

Qy      642  TAGTCTTTTAGCAAAATTTATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATA 701
Db      273993  TATTTTAGAATATATATTAT-TATATATATAATAATAATAATAATAATAATAATAATAATA 274051

Qy      702  AGAAAAATAAATACTACACTAAAGATATACCTACTCTTAGTGAGGCTTATGAAAAAT 761
Db      274052  AATATAGAATATATATTATATATACATATATATAATAATAATAATAATAATAATAATAATA 274111

Qy      762  TTCAAAAAATTTAAATTTTAAAAA-ATAAAAAATTAAGAAAAATGTTTATTACAAGTTGA 820
Db      274112  TATATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 274171

Qy      821  TAAAAAGATCTATTAAAGATTACCTAGTCATATAAGCATTTATTTCAAAGGAAA 872
Db      274172  ATATTTTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 274223
```

Search completed: August 21, 2006, 17:11:49  
Job time : 242 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 21, 2006, 17:11:55 ; Search time 6519 Seconds  
(without alignments)

7855.448 Million cell updates/sec

Title: US-10-734-719-8

Perfect score: 876

Sequence: 1 atgaaaagattattatgc.....attatttcaaggaaataa 876

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 86534536 seqs, 29229259966 residues

Total number of hits satisfying chosen parameters: 173069072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database : Pending Patents NA Main.\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/pna/PCTUSA\_COMB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pna/PCTUSB\_COMB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pna/PCTUSC\_COMB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US075\_COMB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US076\_COMB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US077\_COMB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US078\_COMB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US079\_COMB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US080\_COMB.seq.\*  
10: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US081\_COMB.seq.\*  
11: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US082\_COMB.seq.\*  
12: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US083\_COMB.seq.\*  
13: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US084\_COMB.seq.\*  
14: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US085\_COMB.seq.\*  
15: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US086\_COMB.seq.\*  
16: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US087\_COMB.seq.\*  
17: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US088\_COMB.seq.\*  
18: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US089\_COMB.seq.\*  
19: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US090\_COMB.seq.\*  
20: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US091\_COMB.seq.\*  
21: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US092\_COMB.seq.\*  
22: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US093\_COMB.seq.\*  
23: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US094\_COMB.seq.\*  
24: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US095A\_COMB.seq.\*  
25: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US095B\_COMB.seq.\*  
26: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US095C\_COMB.seq.\*  
27: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US096A\_COMB.seq.\*  
28: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US096B\_COMB.seq.\*  
29: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US096C\_COMB.seq.\*  
30: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US097A\_COMB.seq.\*  
31: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US097B\_COMB.seq.\*  
32: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US098A\_COMB.seq.\*  
33: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US098B\_COMB.seq.\*  
34: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US099A\_COMB.seq.\*  
35: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US099B\_COMB.seq.\*  
36: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US099C\_COMB.seq.\*  
37: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US099D\_COMB.seq.\*  
38: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US099E\_COMB.seq.\*  
39: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US100A\_COMB.seq.\*  
40: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US100B\_COMB.seq.\*  
41: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US101\_COMB.seq.\*  
42: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US102A\_COMB.seq.\*  
43: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US102B\_COMB.seq.\*

44: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US103A\_COMB.seq.\*  
45: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US103B\_COMB.seq.\*  
46: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US103C\_COMB.seq.\*  
47: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US103D\_COMB.seq.\*  
48: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US103E\_COMB.seq.\*  
49: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US103F\_COMB.seq.\*  
50: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US104\_COMB.seq.\*  
51: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US105\_COMB.seq.\*  
52: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US106A\_COMB.seq.\*  
53: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US106B\_COMB.seq.\*  
54: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107A\_COMB.seq.\*  
55: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107B\_COMB.seq.\*  
56: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107C\_COMB.seq.\*  
57: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107D\_COMB.seq.\*  
58: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107E\_COMB.seq.\*  
59: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107F\_COMB.seq.\*  
60: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US107G\_COMB.seq.\*  
61: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US108\_COMB.seq.\*  
62: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US109A\_COMB.seq.\*  
63: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US109B\_COMB.seq.\*  
64: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US109C\_COMB.seq.\*  
65: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US110A\_COMB.seq.\*  
66: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US110B\_COMB.seq.\*  
67: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US110C\_COMB.seq.\*  
68: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US110D\_COMB.seq.\*  
69: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US111A\_COMB.seq.\*  
70: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US111B\_COMB.seq.\*  
71: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US112A\_COMB.seq.\*  
72: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US112B\_COMB.seq.\*  
73: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US113A\_COMB.seq.\*  
74: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US113B\_COMB.seq.\*  
75: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US114\_COMB.seq.\*  
76: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US117\_COMB.seq.\*  
77: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US600\_COMB.seq.\*  
78: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US601\_COMB.seq.\*  
79: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US602A\_COMB.seq.\*  
80: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US602B\_COMB.seq.\*  
81: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US603\_COMB.seq.\*  
82: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US604A\_COMB.seq.\*  
83: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US604B\_COMB.seq.\*  
84: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US605\_COMB.seq.\*  
85: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US606\_COMB.seq.\*  
86: /EMC\_Celerra\_SIDS3/ptodata/2/pna/US607\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	876	100.0	876	47	US-10-303-161-8	Sequence 8, Appl
2	876	100.0	876	60	US-10-734-719-8	Sequence 8, Appl
3	876	100.0	876	61	US-10-820-536-8	Sequence 8, Appl
4	876	100.0	876	61	US-10-821-573-8	Sequence 8, Appl
5	876	100.0	876	61	US-10-821-604-8	Sequence 8, Appl
6	876	100.0	876	61	US-10-830-825-8	Sequence 8, Appl
7	876	100.0	876	61	US-10-830-997-8	Sequence 8, Appl
8	876	100.0	876	61	US-10-845-408-8	Sequence 8, Appl
9	876	100.0	876	61	US-10-845-412-8	Sequence 8, Appl
10	876	100.0	876	61	US-10-846-219-8	Sequence 8, Appl
11	876	100.0	876	61	US-10-847-983-8	Sequence 8, Appl
12	876	100.0	876	61	US-10-850-125-8	Sequence 8, Appl
13	876	100.0	876	61	US-10-850-807-8	Sequence 8, Appl
14	876	100.0	876	63	US-10-961-882-8	Sequence 8, Appl
15	876	100.0	876	63	US-10-962-235-8	Sequence 8, Appl
16	876	100.0	876	63	US-10-962-334-8	Sequence 8, Appl
17	869.8	99.3	873	47	US-10-303-161-13	Sequence 13, Appl
18	869.8	99.3	873	60	US-10-734-719-13	Sequence 13, Appl
19	869.8	99.3	873	61	US-10-820-536-13	Sequence 13, Appl
20	869.8	99.3	873	61	US-10-821-573-13	Sequence 13, Appl



21	869.8	99.3	873	873	61	US-10-821-604-13	Sequence 13, Appl	94	836	95.4	876	63	US-10-961-882-2	Sequence 2, Appl
22	869.8	99.3	873	873	61	US-10-830-825-13	Sequence 13, Appl	95	836	95.4	876	63	US-10-962-235-2	Sequence 2, Appl
23	869.8	99.3	873	873	61	US-10-830-997-13	Sequence 13, Appl	96	836	95.4	876	63	US-10-962-334-2	Sequence 6, Appl
24	869.8	99.3	873	873	61	US-10-845-408-13	Sequence 13, Appl	97	820	93.6	876	47	US-10-303-161-6	Sequence 6, Appl
25	869.8	99.3	873	873	61	US-10-845-412-13	Sequence 13, Appl	98	820	93.6	876	60	US-10-734-719-6	Sequence 6, Appl
26	869.8	99.3	873	873	61	US-10-846-219-13	Sequence 13, Appl	99	820	93.6	876	61	US-10-820-536-6	Sequence 6, Appl
27	869.8	99.3	873	873	61	US-10-847-983-13	Sequence 13, Appl	100	820	93.6	876	61	US-10-821-573-6	Sequence 6, Appl
28	869.8	99.3	873	873	61	US-10-850-125-13	Sequence 13, Appl	101	820	93.6	876	61	US-10-821-604-6	Sequence 6, Appl
29	869.8	99.3	873	873	61	US-10-850-807-13	Sequence 13, Appl	102	820	93.6	876	61	US-10-830-825-6	Sequence 6, Appl
30	869.8	99.3	873	873	61	US-10-850-807-13	Sequence 13, Appl	103	820	93.6	876	61	US-10-830-997-6	Sequence 6, Appl
31	869.8	99.3	873	873	63	US-10-961-882-13	Sequence 13, Appl	104	820	93.6	876	61	US-10-845-408-6	Sequence 6, Appl
32	869.8	99.3	873	873	63	US-10-962-235-13	Sequence 13, Appl	105	820	93.6	876	61	US-10-845-412-6	Sequence 6, Appl
33	866.6	98.9	873	873	47	US-10-303-161-11	Sequence 11, Appl	106	820	93.6	876	61	US-10-846-219-6	Sequence 6, Appl
34	866.6	98.9	873	873	60	US-10-734-719-11	Sequence 11, Appl	107	820	93.6	876	61	US-10-847-983-6	Sequence 6, Appl
35	866.6	98.9	873	873	61	US-10-820-536-11	Sequence 11, Appl	108	820	93.6	876	61	US-10-850-125-6	Sequence 6, Appl
36	866.6	98.9	873	873	61	US-10-821-573-11	Sequence 11, Appl	109	820	93.6	876	61	US-10-850-807-6	Sequence 6, Appl
37	866.6	98.9	873	873	61	US-10-821-604-11	Sequence 11, Appl	110	820	93.6	876	63	US-10-961-882-6	Sequence 6, Appl
38	866.6	98.9	873	873	61	US-10-830-825-11	Sequence 11, Appl	111	820	93.6	876	63	US-10-962-235-6	Sequence 6, Appl
39	866.6	98.9	873	873	61	US-10-830-997-11	Sequence 11, Appl	112	820	93.6	876	63	US-10-962-334-6	Sequence 6, Appl
40	866.6	98.9	873	873	61	US-10-845-408-11	Sequence 11, Appl	113	311.6	35.6	1293	60	US-10-799-016-1	Sequence 1, Appl
41	866.6	98.9	873	873	61	US-10-845-412-11	Sequence 11, Appl	114	286.6	32.7	1830121	13	US-08-426-787-1	Sequence 1, Appl
42	866.6	98.9	873	873	61	US-10-846-219-11	Sequence 11, Appl	115	286.6	32.7	1830121	48	US-10-329-670-1	Sequence 1, Appl
43	866.6	98.9	873	873	61	US-10-847-983-11	Sequence 11, Appl	116	286.6	32.7	1830121	63	US-10-981-687-1	Sequence 1, Appl
44	866.6	98.9	873	873	61	US-10-850-125-11	Sequence 11, Appl	117	281.8	32.2	981	85	US-60-691-214-1696	Sequence 1696, Ap
45	866.6	98.9	873	873	61	US-10-850-807-11	Sequence 11, Appl	118	281.8	32.2	191996	2	PCT-US04-07001-683	Sequence 683, App
46	866.6	98.9	873	873	63	US-10-961-882-11	Sequence 11, Appl	119	281.8	32.2	191996	60	US-10-795-159-683	Sequence 683, App
47	866.6	98.9	873	873	63	US-10-962-235-11	Sequence 11, Appl	120	281.8	32.2	191996	85	US-60-691-214-683	Sequence 683, App
48	866.6	98.9	873	873	63	US-10-962-334-11	Sequence 11, Appl	121	281.8	32.2	1913428	85	US-60-691-214-772	Sequence 772, App
49	844	96.3	876	876	47	US-10-303-161-4	Sequence 4, Appl	122	265.4	30.3	963	85	US-60-691-214-1185	Sequence 1185, Ap
50	844	96.3	876	876	60	US-10-734-719-4	Sequence 4, Appl	123	265.4	30.3	13379	1	PCT-US03-17092-566	Sequence 566, App
51	844	96.3	876	876	61	US-10-820-536-4	Sequence 4, Appl	124	265.4	30.3	13379	2	PCT-US04-07001-566	Sequence 566, App
52	844	96.3	876	876	61	US-10-821-573-4	Sequence 4, Appl	125	265.4	30.3	13379	60	US-10-795-159-566	Sequence 566, App
53	844	96.3	876	876	61	US-10-821-604-4	Sequence 4, Appl	126	265.4	30.3	13379	82	US-60-453-134-566	Sequence 566, App
54	844	96.3	876	876	61	US-10-830-825-4	Sequence 4, Appl	127	265.4	30.3	13379	85	US-60-691-214-566	Sequence 566, App
55	844	96.3	876	876	61	US-10-830-997-4	Sequence 4, Appl	128	265.4	30.3	417576	2	PCT-US04-07001-684	Sequence 684, App
56	844	96.3	876	876	61	US-10-845-408-4	Sequence 4, Appl	129	265.4	30.3	417576	60	US-10-795-159-684	Sequence 684, App
57	844	96.3	876	876	61	US-10-845-412-4	Sequence 4, Appl	130	265.4	30.3	417576	85	US-60-691-214-684	Sequence 684, App
58	844	96.3	876	876	61	US-10-846-219-4	Sequence 4, Appl	131	260.4	29.7	4277	1	PCT-US03-17092-457	Sequence 457, App
59	844	96.3	876	876	61	US-10-847-983-4	Sequence 4, Appl	132	260.4	29.7	4277	2	PCT-US04-07001-457	Sequence 457, App
60	844	96.3	876	876	61	US-10-850-125-4	Sequence 4, Appl	133	260.4	29.7	4277	60	US-10-795-159-457	Sequence 457, App
61	844	96.3	876	876	61	US-10-850-807-4	Sequence 4, Appl	134	260.4	29.7	4277	82	US-60-453-134-457	Sequence 457, App
62	844	96.3	876	876	63	US-10-961-882-4	Sequence 4, Appl	135	260.4	29.7	4277	85	US-60-691-214-457	Sequence 457, App
63	844	96.3	876	876	63	US-10-962-235-4	Sequence 4, Appl	136	255	29.1	909	50	US-10-491-067A-6	Sequence 6, Appl
64	844	96.3	876	876	63	US-10-962-334-4	Sequence 4, Appl	137	95.8	10.9	3673778	48	US-10-312-841A-1	Sequence 1, Appl
65	837.6	95.6	11474	47	US-10-303-161-1	Sequence 1, Appl	138	94.4	10.8	868	28	US-09-654-617-367906	Sequence 367906,	
66	837.6	95.6	11474	60	US-10-734-719-1	Sequence 1, Appl	139	94.4	10.8	868	28	US-09-684-016-367906	Sequence 367906,	
67	837.6	95.6	11474	61	US-10-820-536-1	Sequence 1, Appl	140	94.4	10.8	868	79	US-60-207-458-32398	Sequence 32398, A	
68	837.6	95.6	11474	61	US-10-821-573-1	Sequence 1, Appl	141	89	10.2	8056	50	US-10-473-126-386	Sequence 386, App	
69	837.6	95.6	11474	61	US-10-821-604-1	Sequence 1, Appl	142	88.8	10.1	8056	50	US-10-473-126-386	Sequence 386, App	
70	837.6	95.6	11474	61	US-10-830-825-1	Sequence 1, Appl	143	86.8	9.9	19095	39	US-10-015-127-1938	Sequence 1938, Ap	
71	837.6	95.6	11474	61	US-10-830-997-1	Sequence 1, Appl	144	86.2	9.8	19087	48	US-10-311-455-766	Sequence 766, App	
72	837.6	95.6	11474	61	US-10-845-408-1	Sequence 1, Appl	145	82.8	9.5	219909	75	US-11-405-322-11	Sequence 11, Appl	
73	837.6	95.6	11474	61	US-10-845-412-1	Sequence 1, Appl	146	82.8	9.5	219909	85	US-60-697-446-29	Sequence 29, Appl	
74	837.6	95.6	11474	61	US-10-846-219-1	Sequence 1, Appl	147	82.8	9.5	219909	85	US-60-697-521-29	Sequence 29, Appl	
75	837.6	95.6	11474	61	US-10-847-983-1	Sequence 1, Appl	148	82.8	9.5	219909	86	US-60-723-602-29	Sequence 29, Appl	
76	837.6	95.6	11474	61	US-10-850-125-1	Sequence 1, Appl	149	82.8	9.5	219909	86	US-60-787-402-29	Sequence 29, Appl	
77	837.6	95.6	11474	61	US-10-850-807-1	Sequence 1, Appl	150	82.8	9.5	222273	86	US-60-704-860-18	Sequence 18, Appl	
78	837.6	95.6	11474	63	US-10-961-882-1	Sequence 1, Appl	151	82.6	9.4	8056	50	US-10-473-126-240	Sequence 240, App	
79	837.6	95.6	11474	63	US-10-962-235-1	Sequence 1, Appl	152	81.4	9.3	810	28	US-09-666-355A-6400	Sequence 6400, Ap	
80	837.6	95.6	11474	63	US-10-962-334-1	Sequence 1, Appl	153	81.4	9.3	810	73	US-11-331-032-6400	Sequence 6400, Ap	
81	836	95.4	876	47	US-10-303-161-2	Sequence 2, Appl	154	81.4	9.3	158001	42	US-10-211-179-11	GENERAL INFORMATI	
82	836	95.4	876	60	US-10-734-719-2	Sequence 2, Appl	155	81.2	9.3	219909	75	US-11-405-322-5	Sequence 5, Appl	
83	836	95.4	876	61	US-10-820-536-2	Sequence 2, Appl	156	81.2	9.3	219909	85	US-60-697-446-11	Sequence 11, Appl	
84	836	95.4	876	61	US-10-821-573-2	Sequence 2, Appl	157	81.2	9.3	219909	85	US-60-697-521-11	Sequence 11, Appl	
85	836	95.4	876	61	US-10-821-604-2	Sequence 2, Appl	158	81.2	9.3	219909	86	US-60-723-602-11	Sequence 11, Appl	
86	836	95.4	876	61	US-10-830-825-2	Sequence 2, Appl	159	81.2	9.3	219909	86	US-60-787-402-11	Sequence 11, Appl	
87	836	95.4	876	61	US-10-830-997-2	Sequence 2, Appl	160	81.2	9.3	222273	86	US-60-704-860-10	Sequence 10, Appl	
88	836	95.4	876	61	US-10-845-408-2	Sequence 2, Appl	161	80.8	9.2	1836	74	US-11-360-355-31136	Sequence 31136, A	
89	836	95.4	876	61	US-10-845-412-2	Sequence 2, Appl	162	80.8	9.2	1836	85	US-60-655-875-31136	Sequence 31136, A	
90	836	95.4	876	61	US-10-846-219-2	Sequence 2, Appl	163	80.6	9.2	1282	42	US-10-266-090-36377	Sequence 36377, A	
91	836	95.4												

c 167	79.4	9.1	1770	78	US-60-183-791-6286	Sequence 6286, Ap	c 240	72	8.2	1051	85	US-60-603-421-33298	Sequence 33298, A
168	78.6	9.0	205044	62	US-10-940-774-15851	Sequence 15851, A	241	71.8	8.2	1326	28	US-09-663-779-8205	Sequence 8205, Ap
169	78.6	9.0	205044	62	US-10-940-774-15852	Sequence 15852, A	242	71.8	8.2	1326	65	US-11-049-404-8205	Sequence 8205, Ap
170	78.6	9.0	205044	62	US-10-940-774-15853	Sequence 15853, A	243	71.8	8.2	43614	64	US-10-990-328-94706	Sequence 94706, A
171	78.6	9.0	205044	63	US-10-940-774A-15851	Sequence 15851, A	244	71.8	8.2	43614	64	US-10-990-328A-94706	Sequence 94706, A
172	78.6	9.0	205044	63	US-10-940-774A-15852	Sequence 15852, A	245	71.8	8.2	43614	64	US-10-918-754-16833	Sequence 16833, A
173	78.6	9.0	205044	63	US-10-940-774A-15853	Sequence 15853, A	246	71.8	8.2	129142	62	US-10-990-328-97145	Sequence 97145, A
174	78.6	9.0	223471	62	US-10-940-774-12387	Sequence 12387, A	247	71.8	8.2	129142	83	US-60-495-114-16833	Sequence 16833, A
175	78.6	9.0	223471	62	US-10-940-774-12724	Sequence 12724, A	248	71.8	8.2	129143	64	US-10-990-328-97145	Sequence 97145, A
176	78.6	9.0	223471	62	US-10-940-774-12725	Sequence 12725, A	249	71.8	8.2	129143	3	PCT-US04-37982-2435	Sequence 2435, Ap
177	78.6	9.0	223471	63	US-10-940-774A-12387	Sequence 12387, A	c 250	71.6	8.2	808	79	US-60-207-458-34134	Sequence 34134, A
178	78.6	9.0	223471	63	US-10-940-774A-12724	Sequence 12724, A	c 251	71.6	8.2	834	30	US-09-733-449-2256	Sequence 2256, Ap
179	78.6	9.0	223471	63	US-10-940-774A-12725	Sequence 12725, A	c 252	71.6	8.2	834	32	US-09-803-110-2256	Sequence 2256, Ap
c 180	78.6	9.0	1108228	83	US-60-466-412-86602	Sequence 86602, A	253	71.6	8.2	1164	62	US-10-266-090-37455	Sequence 37455, A
c 181	78.4	8.9	1748	27	US-09-620-392-14063	Sequence 14063, A	c 254	71.6	8.2	1391915	64	US-10-990-328-94753	Sequence 94753, A
c 182	78.4	8.9	1748	30	US-09-702-134-46187	Sequence 46187, A	c 255	71.6	8.2	1391915	64	US-10-990-328A-94753	Sequence 94753, A
c 183	78.4	8.9	1748	32	US-09-815-264-78501	Sequence 78501, A	c 256	71.6	8.2	1391915	83	US-60-466-412-87958	Sequence 87958, A
c 184	78.4	8.9	1748	78	US-60-144-351-2763	Sequence 2763, Ap	c 257	71.6	8.2	2713762	36	US-09-947-911-229	Sequence 229, App
c 185	78	8.9	8056	50	US-10-473-126-240	Sequence 240, App	c 258	71.4	8.2	12201	85	US-60-632-426-233	Sequence 233, App
c 186	77.6	8.9	1016	79	US-60-207-458-33653	Sequence 33653, A	c 259	71.4	8.2	12201	85	US-60-633-250-233	Sequence 233, App
c 187	77.2	8.8	899	26	US-09-565-306-24647	Sequence 24647, A	c 260	71.4	8.2	12201	85	US-60-662-220-233	Sequence 233, App
c 188	77.2	8.8	899	26	US-09-565-306-24647	Sequence 24647, A	c 261	71.4	8.2	12201	86	US-60-723-054-233	Sequence 233, App
c 189	77.2	8.8	1275	42	US-10-266-090-36380	Sequence 36380, A	c 262	71.4	8.2	12201	86	US-60-723-125-233	Sequence 233, App
c 190	77.2	8.8	1368	39	US-10-015-127-2421	Sequence 2421, Ap	c 263	71.4	8.2	12201	86	US-60-740-736-233	Sequence 330, App
c 191	77	8.8	1391915	64	US-10-990-328-94753	Sequence 94753, A	c 264	71.4	8.2	19380	42	US-10-221-613-330	Sequence 390, App
c 192	77	8.8	1391915	64	US-10-990-328A-94753	Sequence 94753, A	c 265	71.4	8.2	27046	79	US-60-229-511-141	Sequence 141, App
c 193	77	8.8	1391915	83	US-60-466-412-87958	Sequence 87958, A	c 266	71.4	8.2	32768	79	US-60-213-178-402	Sequence 402, App
c 194	76.6	8.7	6465	48	US-10-311-455-958	Sequence 958, App	c 267	71.4	8.2	32768	79	US-60-223-511-142	Sequence 142, App
c 195	76.4	8.7	737	23	US-09-474-435A-57848	Sequence 57848, A	c 268	71.4	8.2	158001	42	US-10-211-179-11	GENERAL INFORMATION
c 196	76	8.7	17131	48	US-10-311-455-1026	Sequence 1026, Ap	c 269	71.4	8.2	3037227	36	US-09-947-911-286	Sequence 286, App
c 197	75.6	8.6	1099	28	US-09-666-355A-6603	Sequence 6603, Ap	c 270	71.2	8.1	973	84	US-60-579-062-34683	Sequence 34683, A
c 198	75.6	8.6	1099	73	US-11-331-032-6603	Sequence 6603, Ap	c 271	71.2	8.1	973	85	US-60-603-421-34683	Sequence 34683, A
c 199	75.6	8.6	1501	50	US-10-473-126-328	Sequence 328, App	c 272	71.2	8.1	1063	42	US-10-266-090-10455	Sequence 10455, A
c 200	75.6	8.6	1501	50	US-10-473-126-328	Sequence 328, App	c 273	71.2	8.1	17738	48	US-10-311-455-1512	Sequence 1512, Ap
c 201	75	8.6	1263	42	US-10-266-090-27804	Sequence 27804, A	c 274	71	8.1	269838	3	PCT-US04-37982-3285	Sequence 3285, Ap
c 202	75	8.6	1599662	36	US-09-947-911-108	Sequence 108, App	c 275	70.8	8.1	908	28	US-09-666-355A-6574	Sequence 6574, Ap
c 203	74.8	8.5	2865598	36	US-09-947-911-26	Sequence 26, Appl	c 276	70.8	8.1	908	73	US-11-331-032-6574	Sequence 6574, Ap
c 204	74.4	8.5	1221	42	US-10-266-090-36831	Sequence 36831, A	c 277	70.8	8.1	1164	42	US-10-266-090-32353	Sequence 32353, A
c 205	74	8.4	18988	42	US-10-221-714A-64	Sequence 64, Appl	c 278	70.8	8.1	1372	42	US-10-266-090-21859	Sequence 21859, A
c 206	74	8.4	18988	42	US-10-221-714B-64	Sequence 64, Appl	c 279	70.8	8.1	226706	32	US-09-855-768-1	Sequence 1, Appl
c 207	74	8.4	18988	42	US-10-240-485-62	Sequence 62, Appl	c 280	70.6	8.1	38678	36	US-09-948-947-136	Sequence 136, App
c 208	74	8.4	18988	48	US-10-311-455-674	Sequence 674, App	c 281	70.6	8.1	38678	61	US-10-893-315-136	Sequence 136, App
c 209	73.8	8.4	5930	48	US-10-311-455-490	Sequence 490, App	c 282	70.6	8.1	38684	36	US-09-948-947-154	Sequence 154, App
c 210	73.4	8.4	12592	42	US-10-221-613-58	Sequence 58, Appl	c 283	70.6	8.1	38684	61	US-10-893-315-154	Sequence 154, App
c 211	73.2	8.4	28818	80	US-60-248-542-36	Sequence 26, Appl	c 284	70.6	8.1	1835834	3	PCT-US06-08981-72	Sequence 72, Appl
c 212	73.2	8.4	30808	80	US-60-245-227-30	Sequence 30, Appl	c 285	70.6	8.1	1835834	70	US-11-184-236-72	Sequence 72, Appl
c 213	73.2	8.4	72786	80	US-60-248-498-79	Sequence 79, Appl	c 286	70.6	8.1	3673778	48	US-10-312-841A-1	Sequence 1, Appl
c 214	73.2	8.4	141976	62	US-10-932-349-19364	Sequence 19364, A	c 287	70.4	8.0	537	39	US-10-021-323-8120	Sequence 8120, Ap
c 215	73.2	8.4	141976	84	US-60-500-337-19364	Sequence 19364, A	c 288	70.4	8.0	537	71	US-11-292-078-8120	Sequence 8120, Ap
c 216	73.2	8.4	216452	80	US-60-248-498-17	Sequence 17, Appl	c 289	70.4	8.0	537	80	US-60-255-619-8120	Sequence 8120, Ap
c 217	73	8.3	948	23	US-09-404-520B-13069	Sequence 13069, A	c 290	70.4	8.0	1001	42	US-10-266-090-28485	Sequence 28485, A
c 218	73	8.3	948	23	US-09-404-520B-13069	Sequence 13069, A	c 291	70.4	8.0	193689	64	US-10-990-328-93508	Sequence 93508, A
c 219	72.8	8.3	3673778	48	US-10-312-841A-2	Sequence 2, Appl	c 292	70.4	8.0	193689	64	US-10-990-328A-93508	Sequence 93508, A
c 220	72.6	8.3	824	79	US-60-207-458-18907	Sequence 18907, A	c 293	70.4	8.0	198522	39	US-10-087-192-244	Sequence 244, App
c 221	72.6	8.3	5689	42	US-10-221-714A-148	Sequence 148, App	c 294	70.4	8.0	2307596	36	US-09-948-128-334	Sequence 334, App
c 222	72.6	8.3	5689	42	US-10-221-714B-148	Sequence 148, App	c 295	70.2	8.0	1143	79	US-60-207-458-36757	Sequence 36757, A
c 223	72.6	8.3	5689	42	US-10-239-676-90	Sequence 90, Appl	c 296	70.2	8.0	8201	42	US-10-240-452-27	Sequence 27, Appl
c 224	72.6	8.3	5689	42	US-10-240-453-100	Sequence 100, App	c 297	70.2	8.0	8201	48	US-10-311-455-279	Sequence 279, App
c 225	72.4	8.3	865	84	US-60-579-062-25190	Sequence 25190, A	c 298	70.2	8.0	13286	51	US-10-517-441-746	Sequence 746, App
c 226	72.4	8.3	865	85	US-60-603-421-25190	Sequence 25190, A	c 299	70	8.0	420	32	US-09-804-730-10807	Sequence 10807, A
c 227	72.4	8.3	980	28	US-09-666-355A-6399	Sequence 6399, Ap	c 300	70	8.0	420	78	US-60-189-657-11056	Sequence 11056, A
c 228	72.4	8.3	980	73	US-11-331-032-6399	Sequence 6399, Ap	c 301	70	8.0	6314	42	US-10-240-452-14	Sequence 14, Appl
c 229	72.4	8.3	988	28	US-09-666-355A-774	Sequence 774, App	c 302	70	8.0	98950	3	PCT-US05-10912-4	Sequence 4, Appl
c 230	72.4	8.3	988	73	US-11-331-032-774	Sequence 774, App	c 303	69.8	8.0	627	23	US-09-474-435A-45851	Sequence 45851, A
c 231	72.4	8.3	5979	42	US-10-239-676-18	Sequence 18, Appl	c 304	69.8	8.0	821	79	US-60-207-458-33231	Sequence 33231, A
c 232	72.4	8.3	5979	42	US-10-240-453-26	Sequence 26, Appl	c 305	69.8	8.0	1090	42	US-10-266-090-12176	Sequence 12176, A
c 233	72.4	8.3	6132	48	US-10-311-455-836	Sequence 836, App	c 306	69.8	8.0	12201	85	US-60-632-426-105	Sequence 105, App
c 234	72.4	8.3	1317241	3	PCT-US06-08981-73	Sequence 73, Appl	c 307	69.8	8.0	12201	85	US-60-633-250-105	Sequence 105, App
c 235	72.4	8.3	1317241	70	US-11-184-236-73	Sequence 73, Appl	c 308	69.8	8.0	12201	85	US-60-662-220-105	Sequence 105, App
c 236	72.2	8.2	49298	83	PCT-US06-412-85631	Sequence 85631, A	c 309	69.8	8.0	12201	86	US-60-723-054-105	Sequence 105, App
c 237	72.2	8.2	1971884	3	PCT-US06-08981-70	Sequence 70, Appl	c 310	69.8	8.0	12201	86	US-60-723-125-105	Sequence 105, App
c 238	72.2	8.2	1971884	70	US-11-184-236-70	Sequence 70, Appl	c 311	69.8	8.0	12201	86	US-60-740-736-105	Sequence 105, App
c 239	72	8.2	1051	84	US-60-579-062-33298	Sequence 33298, A	c 312	69.8	8.0	13131	42	US-10-240-599C-58	Sequence 58, Appl

c 313	69.8	8.0	28054	79	US-60-212-657-88	Sequence 88, Appl	386	68	7.8	396	26	US-09-565-306-57274	Sequence 67274, A
c 314	69.8	8.0	32768	79	US-60-213-181-33	Sequence 33, Appl	387	68	7.8	396	28	US-09-654-617-330650	Sequence 330650, A
c 315	69.8	8.0	32768	79	US-60-213-181-34	Sequence 34, Appl	388	68	7.8	396	28	US-09-684-016-330650	Sequence 330650, A
c 316	69.8	8.0	32768	79	US-60-213-181-35	Sequence 35, Appl	389	68	7.8	396	74	US-11-353-150-57274	Sequence 67274, A
c 317	69.8	8.0	1317241	3	PCT-US06-08981-73	Sequence 73, Appl	390	68	7.8	487	28	US-09-654-617-336985	Sequence 336985, A
c 318	69.8	8.0	1317241	3	US-11-184-236-73	Sequence 73, Appl	391	68	7.8	487	28	US-09-684-016-336985	Sequence 336985, A
c 319	69.8	8.0	2111522	36	US-09-947-911-63	Sequence 63, Appl	392	68	7.8	487	29	US-09-652-257A-354	Sequence 354, App
c 320	69.6	7.9	927	85	US-60-579-062-39891	Sequence 39891, A	393	68	7.8	487	78	US-60-162-747-375	Sequence 375, App
c 321	69.6	7.9	927	85	US-60-579-062-39891	Sequence 39891, A	394	68	7.8	539	23	US-09-474-435A-39400	Sequence 39400, A
c 322	69.6	7.9	927	85	US-60-579-062-39891	Sequence 39891, A	395	68	7.8	539	23	US-09-474-435A-39400	Sequence 39400, A
c 323	69.6	7.9	1013	84	US-60-579-062-31292	Sequence 31292, A	396	68	7.8	573	33	US-09-874-708A-9961	Sequence 9961, Ap
c 324	69.6	7.9	1013	84	US-60-579-062-31292	Sequence 31292, A	397	68	7.8	573	33	US-09-874-708A-9961	Sequence 9961, Ap
c 325	69.6	7.9	34196	80	US-60-603-421-31292	Sequence 31292, A	c 398	68	7.8	644	26	US-09-565-306-26090	Sequence 26090, A
c 326	69.6	7.9	43515	82	US-60-245-228-189	Sequence 189, App	c 399	68	7.8	644	26	US-09-565-306-26090	Sequence 26090, A
c 327	69.6	7.9	43515	82	US-60-465-241-52566	Sequence 52566, A	c 400	68	7.8	858	84	US-60-579-062-22005	Sequence 22005, A
c 328	69.6	7.9	43515	82	US-60-465-241-52566	Sequence 52566, A	c 401	68	7.8	858	84	US-60-579-062-22005	Sequence 22005, A
c 329	69.4	7.9	867	70	US-09-666-355A-6392	Sequence 6392, Ap	c 402	68	7.8	7903	42	US-10-239-676-110	Sequence 110, App
c 330	69.4	7.9	867	70	US-10-767-795-55437	Sequence 55437, A	c 403	68	7.8	7903	42	US-10-240-453-122	Sequence 122, App
c 331	69.4	7.9	867	70	US-11-331-032-6392	Sequence 6392, Ap	c 404	68	7.8	7903	48	US-10-311-455-1336	Sequence 1336, Ap
c 332	69.4	7.9	2540030	3	PCT-US06-08981-69	Sequence 69, Appl	c 405	68	7.8	1000000	47	US-10-308-107-2	Sequence 2, Appli
c 333	69.4	7.9	2540030	3	US-11-184-236-69	Sequence 69, Appl	406	67.8	7.7	335	27	US-09-619-643-12197	Sequence 12197, A
c 334	69.2	7.9	610	33	US-09-874-708A-7626	Sequence 7626, Ap	407	67.8	7.7	335	28	US-09-654-617-340051	Sequence 340051, A
c 335	69.2	7.9	610	33	US-09-874-708A-7626	Sequence 7626, Ap	408	67.8	7.7	335	28	US-09-654-617-340051	Sequence 340051, A
c 336	69.2	7.9	860	32	US-09-824-557-79	Sequence 79, Appl	409	67.8	7.7	335	71	US-11-227-183A-12197	Sequence 12197, A
c 337	69.2	7.9	7892	42	US-10-257-166-138	Sequence 138, App	c 410	67.8	7.7	335	78	US-60-145-485-6478	Sequence 6478, Ap
c 338	69.2	7.9	7892	42	US-10-257-166-138	Sequence 138, App	c 411	67.8	7.7	1145	42	US-10-266-090-23096	Sequence 23096, A
c 339	69.2	7.9	38972	64	US-10-990-328-95656	Sequence 95656, A	c 412	67.8	7.7	1145	42	US-10-266-090-23096	Sequence 23096, A
c 340	69.2	7.9	38972	64	US-10-990-328-95656	Sequence 95656, A	c 413	67.8	7.7	6641	42	US-60-762-056-27923	Sequence 27923, A
c 341	69	7.9	187169	62	US-10-940-774-12776	Sequence 12776, A	c 414	67.8	7.7	6641	48	US-10-240-452-35	Sequence 35, Appl
c 342	69	7.9	187169	62	US-10-940-774-12776	Sequence 12776, A	c 415	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 343	69	7.9	191569	63	US-10-940-774-15940	Sequence 15940, A	c 416	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 344	68.8	7.9	565	33	US-09-874-708A-34471	Sequence 34471, A	c 417	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 345	68.8	7.9	565	33	US-09-874-708A-34471	Sequence 34471, A	c 418	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 346	68.8	7.9	932	32	US-09-849-526A-20468	Sequence 20468, A	c 419	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 347	68.8	7.9	932	32	US-09-849-526A-20468	Sequence 20468, A	c 420	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 348	68.8	7.9	25590	36	US-09-949-003C-5007	Sequence 5007, Ap	c 421	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 349	68.8	7.9	25590	36	US-11-033-545-777	Sequence 777, App	c 422	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 350	68.8	7.9	2347870	36	US-09-947-911-294	Sequence 294, Appl	c 423	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 351	68.8	7.9	13831263	36	US-09-947-911-294	Sequence 294, Appl	c 424	67.8	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 352	68.6	7.8	777	28	US-09-666-355A-1045	Sequence 1045, Ap	c 425	67.6	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 353	68.6	7.8	777	28	US-10-311-032-1045	Sequence 1045, Ap	c 426	67.6	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 354	68.6	7.8	17967	48	US-10-311-455-987	Sequence 987, App	c 427	67.6	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 355	68.4	7.8	897	84	US-60-579-062-37549	Sequence 37549, A	c 428	67.6	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 356	68.4	7.8	897	84	US-60-579-062-37549	Sequence 37549, A	c 429	67.6	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 357	68.4	7.8	1279	42	US-10-266-090-28045	Sequence 28045, A	c 430	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 358	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 431	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 359	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 432	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 360	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 433	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 361	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 434	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 362	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 435	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 363	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 436	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 364	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 437	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 365	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 438	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 366	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 439	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 367	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 440	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 368	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 441	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 369	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 442	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 370	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 443	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 371	68.4	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 444	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 372	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 445	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 373	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 446	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 374	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 447	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 375	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 448	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 376	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 449	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 377	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 450	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 378	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 451	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 379	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 452	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 380	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 453	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 381	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 454	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 382	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 455	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 383	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 456	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 384	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 457	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App
c 385	68.2	7.8	4633	27	US-09-614-150-17206	Sequence 17206, A	c 458	67.4	7.7	7932	48	US-10-311-455-287	Sequence 287, App

C 459	67.2	7.7	1501	50	US-10-496-475-158	Sequence 158, App	C 532	66.6	7.6	12138	42	US-10-257-166B-116	Sequence 116, App
C 460	67.2	7.7	8170	42	US-10-240-453-132	Sequence 132, App	C 533	66.6	7.6	12138	48	US-10-311-455-1602	Sequence 1602, App
461	67	7.6	676	32	US-09-865-419A-7909	Sequence 7909, App	C 534	66.6	7.6	12138	48	US-10-311-455-1602	Sequence 1915, App
462	67	7.6	745	23	US-09-474-436-352	Sequence 352, App	535	66.6	7.6	18032	62	US-10-932-349-19145	Sequence 19145, A
463	67	7.6	745	28	US-09-637-886-224	Sequence 224, App	536	66.6	7.6	18032	84	US-60-500-337-19145	Sequence 19145, A
464	67	7.6	745	28	US-09-652-911-1505	Sequence 1505, App	537	66.6	7.6	30810	36	US-09-949-003C-4357	Sequence 4357, App
465	67	7.6	805	78	US-60-126-265-127	Sequence 127, App	538	66.6	7.6	30967	36	US-09-949-003C-5154	Sequence 5154, App
466	67	7.6	828	79	US-60-207-458-58087	Sequence 58087, A	539	66.6	7.6	94709	64	US-10-990-328-93703	Sequence 93703, A
467	67	7.6	929	28	US-09-669-817A-31403	Sequence 31403, A	540	66.6	7.6	94709	64	US-10-990-328-93703	Sequence 93703, A
C 468	67	7.6	1003	42	US-10-266-090-37449	Sequence 37449, A	C 541	66.6	7.6	213862	64	US-10-990-328-97913	Sequence 97913, A
C 469	67	7.6	1008	42	US-10-266-090-30421	Sequence 30421, A	C 542	66.6	7.6	213862	64	US-10-990-328A-97913	Sequence 97913, A
C 470	67	7.6	1018	42	US-10-266-090-27825	Sequence 27825, A	C 543	66.4	7.6	371	22	US-09-304-517A-267498	Sequence 267498, A
C 471	67	7.6	1035	42	US-10-266-090-34486	Sequence 34486, A	C 544	66.4	7.6	371	22	US-09-371-146A-266361	Sequence 266361, A
C 472	67	7.6	2411	74	US-11-360-355-25925	Sequence 25925, A	C 545	66.4	7.6	371	26	US-09-565-240-22039	Sequence 22039, A
C 473	67	7.6	2411	85	US-60-655-875-25925	Sequence 25925, A	C 546	66.4	7.6	371	28	US-09-654-617-17151	Sequence 17151, A
C 474	67	7.6	7498	48	US-10-311-455-230	Sequence 230, App	C 547	66.4	7.6	371	28	US-09-684-016-17151	Sequence 17151, A
C 475	67	7.6	8900	51	US-10-517-441-472	Sequence 441, App	C 548	66.4	7.6	371	38	US-09-985-678-267498	Sequence 267498, A
C 476	67	7.6	8900	51	US-10-517-441-702	Sequence 702, App	C 549	66.4	7.6	371	71	US-11-239-610A-22039	Sequence 22039, A
C 477	67	7.6	13286	51	US-10-517-441-472	Sequence 472, App	C 550	66.4	7.6	601	62	US-10-940-774-30531	Sequence 30531, A
C 478	67	7.6	16287	48	US-10-311-455-646	Sequence 646, App	C 551	66.4	7.6	601	62	US-10-940-774-37150	Sequence 37150, A
C 479	67	7.6	16885	83	US-60-466-412-85418	Sequence 85418, A	C 552	66.4	7.6	601	62	US-10-940-774-37164	Sequence 37164, A
C 480	67	7.6	27978	79	US-60-207-423-22	Sequence 22, Appl	C 553	66.4	7.6	601	62	US-10-940-774-145868	Sequence 145868, A
C 481	67	7.6	32768	79	US-60-207-423-23	Sequence 23, Appl	C 554	66.4	7.6	601	62	US-10-940-774-146136	Sequence 146136, A
C 482	67	7.6	32768	79	US-60-207-905-22	Sequence 22, Appl	C 555	66.4	7.6	601	63	US-10-940-774-146404	Sequence 146404, A
C 483	67	7.6	32768	79	US-60-207-905-23	Sequence 23, Appl	C 556	66.4	7.6	601	63	US-10-940-774A-30531	Sequence 30531, A
C 484	67	7.6	32768	79	US-60-207-905-24	Sequence 24, Appl	C 557	66.4	7.6	601	63	US-10-940-774A-37150	Sequence 37150, A
C 485	67	7.6	32768	79	US-60-208-837-26	Sequence 26, Appl	C 558	66.4	7.6	601	63	US-10-940-774A-37164	Sequence 37164, A
C 486	67	7.6	32768	79	US-60-208-837-27	Sequence 27, Appl	C 559	66.4	7.6	601	63	US-10-940-774A-145868	Sequence 145868, A
C 487	67	7.6	32768	79	US-60-208-837-28	Sequence 28, Appl	C 560	66.4	7.6	601	63	US-10-940-774A-146136	Sequence 146136, A
C 488	67	7.6	360697	3	PCT-US04-37982-2113	Sequence 2113, Ap	C 561	66.4	7.6	601	63	US-10-940-774A-146404	Sequence 146404, A
489	66.8	7.6	430	28	US-09-654-617-348605	Sequence 348605, A	C 562	66.4	7.6	721	49	US-10-363-345A-3373	Sequence 3373, Ap
490	66.8	7.6	430	28	US-09-684-016-348605	Sequence 348605, A	C 563	66.4	7.6	721	49	US-10-363-345A-3373	Sequence 3374, Ap
491	66.8	7.6	430	29	US-09-684-016-348605	Sequence 348605, A	C 564	66.4	7.6	721	49	US-10-363-483A-3373	Sequence 3373, Ap
492	66.8	7.6	430	73	US-11-330-364-371	Sequence 371, App	C 565	66.4	7.6	721	49	US-10-363-483A-3374	Sequence 3374, Ap
493	66.8	7.6	430	78	US-60-161-619-371	Sequence 371, App	C 566	66.4	7.6	754	79	US-60-207-458-288	Sequence 288, App
C 494	66.8	7.6	733	33	US-09-865-439A-48346	Sequence 48346, A	C 567	66.4	7.6	821	84	US-60-579-062-36688	Sequence 36688, A
C 495	66.8	7.6	733	33	US-09-865-439A-48346	Sequence 92609, A	C 568	66.4	7.6	821	85	US-60-603-421-36688	Sequence 36688, A
C 496	66.8	7.6	799	23	US-09-404-520-13067	Sequence 13067, A	C 569	66.4	7.6	856	84	US-60-579-062-22430	Sequence 22430, A
C 497	66.8	7.6	799	23	US-09-404-520B-13067	Sequence 13067, A	C 570	66.4	7.6	856	85	US-60-603-421-22430	Sequence 22430, A
C 498	66.8	7.6	897	84	US-60-579-062-36113	Sequence 36113, A	C 571	66.4	7.6	1029	50	US-10-425-115-123086	Sequence 123086, A
C 499	66.8	7.6	897	85	US-60-603-421-36113	Sequence 36113, A	C 572	66.4	7.6	1153	42	US-10-266-090-29892	Sequence 29892, A
C 500	66.8	7.6	979	45	US-10-301-480A-238509	Sequence 238509, A	C 573	66.4	7.6	2300	86	US-60-762-056-35494	Sequence 35494, A
501	66.8	7.6	979	45	US-10-301-480A-238510	Sequence 238510, A	C 574	66.4	7.6	32768	79	US-60-212-413-130	Sequence 130, App
502	66.8	7.6	979	45	US-10-301-480A-238511	Sequence 238511, A	C 575	66.4	7.6	299970	62	US-10-918-754-16375	Sequence 16375, A
503	66.8	7.6	979	45	US-10-301-480A-238512	Sequence 238512, A	C 576	66.4	7.6	299970	83	US-60-495-114-16375	Sequence 16375, A
504	66.8	7.6	979	45	US-10-301-480A-238516	Sequence 238516, A	C 577	66.2	7.6	401	27	US-09-619-643-22144	Sequence 22144, A
505	66.8	7.6	979	46	US-10-301-480B-238509	Sequence 238509, A	578	66.2	7.6	401	28	US-09-654-617-344474	Sequence 344474, A
506	66.8	7.6	979	46	US-10-301-480B-238510	Sequence 238510, A	579	66.2	7.6	401	28	US-09-684-016-344474	Sequence 344474, A
507	66.8	7.6	979	46	US-10-301-480B-238511	Sequence 238511, A	580	66.2	7.6	401	71	US-11-227-183A-22144	Sequence 22144, A
508	66.8	7.6	979	46	US-10-301-480B-238512	Sequence 238512, A	581	66.2	7.6	401	78	US-60-145-147-1373	Sequence 1373, Ap
509	66.8	7.6	979	46	US-10-301-480B-238516	Sequence 238516, A	582	66.2	7.6	412	28	US-09-654-617-85926	Sequence 85926, A
510	66.8	7.6	979	47	US-10-301-480C-238509	Sequence 238509, A	583	66.2	7.6	412	28	US-09-684-016-85926	Sequence 85926, A
511	66.8	7.6	979	47	US-10-301-480C-238510	Sequence 238510, A	C 584	66.2	7.6	546	36	US-09-948-947-1192	Sequence 1192, Ap
512	66.8	7.6	979	47	US-10-301-480C-238511	Sequence 238511, A	C 585	66.2	7.6	546	61	US-10-893-315-1192	Sequence 1192, Ap
513	66.8	7.6	979	47	US-10-301-480C-238512	Sequence 238512, A	C 586	66.2	7.6	548	36	US-09-948-947-1191	Sequence 1191, Ap
514	66.8	7.6	979	47	US-10-301-480C-238516	Sequence 238516, A	C 587	66.2	7.6	548	61	US-10-893-315-1191	Sequence 1191, Ap
C 515	66.8	7.6	3366	3	PCT-US99-26796-308	Sequence 308, App	C 588	66.2	7.6	577	33	US-09-874-708A-11415	Sequence 11415, A
C 516	66.8	7.6	193689	64	US-10-990-328-93508	Sequence 93508, A	C 589	66.2	7.6	577	79	US-60-211-750-11272	Sequence 11272, A
C 517	66.8	7.6	193689	64	US-10-990-328A-93508	Sequence 93508, A	C 590	66.2	7.6	637	28	US-09-654-617-429291	Sequence 429291, A
C 518	66.8	7.6	198522	39	US-10-087-192-334	Sequence 244, App	C 591	66.2	7.6	637	28	US-09-669-817A-12145	Sequence 12145, A
C 519	66.8	7.6	2307596	36	US-09-948-128-334	Sequence 334, App	C 592	66.2	7.6	637	28	US-09-684-016-429291	Sequence 429291, A
C 520	66.8	7.6	7413081	36	US-09-948-128-399	Sequence 399, App	C 593	66.2	7.6	898	42	US-10-266-090-23655	Sequence 23655, A
C 521	66.6	7.6	793	79	US-60-207-458-18916	Sequence 18916, A	C 594	66.2	7.6	1218	42	US-10-266-090-12154	Sequence 12154, A
C 522	66.6	7.6	884	84	US-60-579-062-15764	Sequence 15764, A	C 595	66.2	7.6	84452	62	US-10-932-333-11428	Sequence 11428, A
C 523	66.6	7.6	884	85	US-60-603-421-15764	Sequence 15764, A	C 596	66.2	7.6	84452	84	US-60-500-315-11428	Sequence 11428, A
C 524	66.6	7.6	1190	79	US-60-207-458-36950	Sequence 36950, A	C 597	66	7.5	431	23	US-09-480-902-5558	Sequence 5558, Ap
C 525	66.6	7.6	2300	86	US-60-762-056-42403	Sequence 42403, A	C 598	66	7.5	431	37	US-09-960-352-5558	Sequence 5558, Ap
C 526	66.6	7.6	5273	48	US-10-311-455-848	Sequence 848, App	C 599	66	7.5	568	32	US-09-849-526A-4828	Sequence 4828, Ap
C 527	66.6	7.6	5893	48	US-10-311-455-832	Sequence 832, App	C 600	66	7.5	568	79	US-60-202-214-4828	Sequence 4828, Ap
C 528	66.6	7.6	5928	48	US-10-311-455-2059	Sequence 2059, App	C 601	66	7.5	787	30	US-09-739-449-2501	Sequence 2501, Ap
C 529	66.6	7.6	6012	42	US-10-221-613-236	Sequence 236, App	C 602	66	7.5	787	32	US-09-803-110-2501	Sequence 2501, Ap
C 530	66.6	7.6	12138	42	US-10-240-453-209	Sequence 209, App	C 603	66	7.5	847	23	US-09-406-292A-285	Sequence 285, App
C 531	66.6	7.6	12138	42	US-10-257-166-116	Sequence 116, App	C 604	66	7.5	979	45	US-10-301-480A-238513	Sequence 238513, A

605	66	7.5	979	45	US-10-301-480A-238515	Sequence 238515, A	c 678	65.4	7.5	466	74	US-11-353-150-29077	Sequence 29077, A
606	66	7.5	979	46	US-10-301-480B-238513	Sequence 238513, A	c 679	65.4	7.5	30810	36	US-09-949-003C-4357	Sequence 4357, Ap
607	66	7.5	979	47	US-10-301-480B-238515	Sequence 238515, A	c 680	65.4	7.5	30967	36	US-09-949-003C-4354	Sequence 4354, Ap
608	66	7.5	979	47	US-10-301-480B-238513	Sequence 238513, A	c 681	65.4	7.5	343684	64	US-10-990-328-95986	Sequence 95986, A
609	66	7.5	979	47	US-10-301-480B-238515	Sequence 238515, A	c 682	65.4	7.5	343684	64	US-10-990-328-95986	Sequence 95986, A
610	66	7.5	980	45	US-10-301-480A-238519	Sequence 238519, A	c 683	65.4	7.5	1835834	3	PCT-US06-08981-72	Sequence 72, Appl
611	66	7.5	980	46	US-10-301-480B-238519	Sequence 238519, A	c 684	65.4	7.5	1835834	70	US-11-184-236-72	Sequence 72, Appl
612	66	7.5	980	47	US-10-301-480C-238519	Sequence 238519, A	c 685	65.2	7.4	489	27	US-09-619-643-25783	Sequence 25783, A
c 613	66	7.5	1052	23	US-09-406-292A-413	Sequence 413, App	686	65.2	7.4	489	27	US-09-619-643-25783	Sequence 25783, A
614	66	7.5	1232	42	US-10-266-090-32307	Sequence 32307, A	687	65.2	7.4	489	28	US-09-654-617-346371	Sequence 346371, A
615	66	7.5	1580	42	US-10-266-090-12229	Sequence 12229, A	688	65.2	7.4	489	28	US-09-654-617-346371	Sequence 346371, A
616	66	7.5	2300	86	US-60-762-056-4574	Sequence 4574, Ap	689	65.2	7.4	489	71	US-11-227-183A-25783	Sequence 25783, A
c 617	66	7.5	22350	80	US-60-245-228-17	Sequence 17, Appl	690	65.2	7.4	489	78	US-09-146-905-1603	Sequence 1603, Ap
618	66	7.5	29300	1	PCT-US01-32004-1	Sequence 1, Appli	c 691	65.2	7.4	666	33	US-09-865-439A-44547	Sequence 44547, A
c 619	66	7.5	38972	64	US-10-990-328-95856	Sequence 95856, A	692	65.2	7.4	666	79	US-60-207-458-88810	Sequence 88810, A
c 620	66	7.5	38972	64	US-10-990-328A-95856	Sequence 95856, A	693	65.2	7.4	980	45	US-10-301-480A-238517	Sequence 238517, A
c 621	66	7.5	125076	36	US-09-947-916-29	Sequence 29, Appl	694	65.2	7.4	980	45	US-10-301-480A-238518	Sequence 238518, A
c 622	66	7.5	162075	79	US-60-226-176-1063	Sequence 1063, Ap	695	65.2	7.4	980	46	US-10-301-480B-238517	Sequence 238517, A
c 623	66	7.5	162075	79	US-60-233-468-1063	Sequence 1063, Ap	696	65.2	7.4	980	46	US-10-301-480B-238518	Sequence 238518, A
c 624	66	7.5	162075	81	US-60-313-371-1063	Sequence 1063, Ap	697	65.2	7.4	980	47	US-10-301-480C-238517	Sequence 238517, A
c 625	66	7.5	210071	80	US-60-242-679-302	Sequence 302, App	698	65.2	7.4	980	47	US-10-301-480C-238518	Sequence 238518, A
626	65.8	7.5	413	28	US-09-637-086A-29308	Sequence 29308, A	c 699	65.2	7.4	5739	48	US-10-266-090-23096	Sequence 23096, A
627	65.8	7.5	413	28	US-09-637-086B-29308	Sequence 29308, A	c 700	65.2	7.4	5739	48	US-10-311-455-691	Sequence 691, App
628	65.8	7.5	413	73	US-11-331-019-29308	Sequence 29308, A	c 701	65.2	7.4	5739	48	US-10-221-714A-165	Sequence 165, App
c 629	65.8	7.5	434	28	US-09-654-617-425761	Sequence 425761, A	c 702	65.2	7.4	5739	48	US-10-221-714B-165	Sequence 165, App
c 630	65.8	7.5	434	28	US-09-669-817A-28236	Sequence 28236, A	c 703	65.2	7.4	5739	48	US-10-240-589C-49	Sequence 49, Appl
c 631	65.8	7.5	434	28	US-09-669-817A-28236	Sequence 28236, A	c 704	65.2	7.4	5739	48	US-10-240-589C-49	Sequence 49, Appl
c 632	65.8	7.5	673	50	US-10-424-599-57885	Sequence 425761, A	c 705	65.2	7.4	5739	48	US-10-506-111-199	Sequence 199, App
c 633	65.8	7.5	731	84	US-60-579-062-37038	Sequence 37038, A	c 706	65.2	7.4	5739	48	US-10-311-455-2405	Sequence 2405, Ap
c 634	65.8	7.5	731	85	US-60-603-421-37038	Sequence 37038, A	c 707	65.2	7.4	5739	48	US-10-257-166B-1	Sequence 1, Appli
c 635	65.8	7.5	979	45	US-10-301-480A-238508	Sequence 238508, A	c 708	65.2	7.4	5739	48	US-10-311-455-77	Sequence 77, Appl
c 636	65.8	7.5	979	46	US-10-301-480B-238508	Sequence 238508, A	c 709	65.2	7.4	5739	48	US-10-311-455-77	Sequence 77, Appl
c 637	65.8	7.5	979	47	US-10-301-480C-238508	Sequence 238508, A	c 710	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 638	65.8	7.5	1111	79	US-60-207-458-37399	Sequence 37399, A	c 711	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 639	65.8	7.5	1111	42	US-10-266-090-20483	Sequence 20483, A	c 712	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 640	65.8	7.5	7442	42	US-10-221-714A-409	Sequence 409, App	c 713	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 641	65.8	7.5	7442	42	US-10-221-714B-409	Sequence 409, App	c 714	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 642	65.8	7.5	7544	42	US-10-239-676-5	Sequence 5, Appli	c 715	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 643	65.8	7.5	7544	42	US-10-240-453-13	Sequence 13, Appl	c 716	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 644	65.8	7.5	9964	48	US-10-311-455-71	Sequence 71, Appl	c 717	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 645	65.6	7.5	516	23	US-09-480-352-5785	Sequence 5785, Ap	c 718	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 646	65.6	7.5	516	23	US-09-480-352-5785	Sequence 5785, Ap	c 719	65.2	7.4	5739	48	US-10-940-774A-16110	Sequence 16110, A
c 647	65.6	7.5	737	23	US-09-474-435A-57848	Sequence 57848, A	c 720	65.2	7.4	5739	48	US-10-221-714A-266	Sequence 266, App
c 648	65.6	7.5	963	42	US-10-266-090-10200	Sequence 10200, A	c 721	65.2	7.4	5739	48	US-10-221-714B-266	Sequence 266, App
c 649	65.6	7.5	979	44	US-10-301-480-595013	Sequence 595013, A	c 722	65.2	7.4	5739	48	PCT-US01-32004-1	Sequence 1, Appli
c 650	65.6	7.5	979	44	US-10-301-480-595014	Sequence 595014, A	c 723	65.2	7.4	5739	48	US-10-918-754-16833	Sequence 16833, A
c 651	65.6	7.5	979	44	US-10-301-480-595015	Sequence 595015, A	c 724	65.2	7.4	5739	48	US-60-495-114-16833	Sequence 16833, A
c 652	65.6	7.5	979	44	US-10-301-480-1208422	Sequence 1208422, A	c 725	65.2	7.4	5739	48	US-60-495-114-16833	Sequence 16833, A
c 653	65.6	7.5	979	44	US-10-301-480-1208423	Sequence 1208423, A	c 726	65.2	7.4	5739	48	US-10-990-328-97145	Sequence 97145, A
c 654	65.6	7.5	979	44	US-10-301-480-1208424	Sequence 1208424, A	c 727	65.2	7.4	5739	48	US-10-990-328A-97145	Sequence 97145, A
c 655	65.6	7.5	979	45	US-10-301-480A-238514	Sequence 238514, A	c 728	65.2	7.4	5739	48	PCT-US04-37982-2435	Sequence 2435, Ap
c 656	65.6	7.5	979	45	US-10-301-480A-238514	Sequence 238514, A	c 729	65.2	7.4	5739	48	US-60-226-176-1063	Sequence 1063, Ap
c 657	65.6	7.5	979	45	US-10-301-480A-971082	Sequence 971082, A	c 730	65.2	7.4	5739	48	US-60-233-468-1063	Sequence 1063, Ap
c 658	65.6	7.5	979	45	US-10-301-480A-971083	Sequence 971083, A	c 731	65.2	7.4	5739	48	US-60-313-371-1063	Sequence 1063, Ap
c 659	65.6	7.5	979	46	US-10-301-480B-238514	Sequence 238514, A	c 732	65.2	7.4	5739	48	US-10-308-107-2	Sequence 2, Appli
c 660	65.6	7.5	979	46	US-10-301-480B-971082	Sequence 971082, A	c 733	65.2	7.4	5739	48	US-09-654-617-428779	Sequence 428779, A
c 661	65.6	7.5	979	46	US-10-301-480B-971083	Sequence 971083, A	c 734	65.2	7.4	5739	48	US-09-654-617-428779	Sequence 428779, A
c 662	65.6	7.5	979	46	US-10-301-480B-971084	Sequence 971084, A	c 735	65.2	7.4	5739	48	US-09-654-617-428779	Sequence 428779, A
c 663	65.6	7.5	979	47	US-10-301-480C-238514	Sequence 238514, A	c 736	65.2	7.4	5739	48	US-09-304-517A-275714	Sequence 275714, A
c 664	65.6	7.5	979	47	US-10-301-480C-971082	Sequence 971082, A	c 737	65.2	7.4	5739	48	US-09-371-146A-274495	Sequence 274495, A
c 665	65.6	7.5	979	47	US-10-301-480C-971083	Sequence 971083, A	c 738	65.2	7.4	5739	48	US-09-440-687-28512	Sequence 28512, A
c 666	65.6	7.5	979	47	US-10-301-480C-971084	Sequence 971084, A	c 739	65.2	7.4	5739	48	US-09-440-687-28512	Sequence 28512, A
c 667	65.6	7.5	980	44	US-10-301-480-595016	Sequence 595016, A	c 740	65.2	7.4	5739	48	US-09-615-606A-52918	Sequence 52918, A
c 668	65.6	7.5	980	44	US-10-301-480-1208425	Sequence 1208425, A	c 741	65.2	7.4	5739	48	US-09-615-606A-52918	Sequence 52918, A
c 669	65.6	7.5	980	45	US-10-301-480B-971085	Sequence 971085, A	c 742	65.2	7.4	5739	48	US-09-619-643-3569	Sequence 3569, Ap
c 670	65.6	7.5	980	46	US-10-301-480B-971085	Sequence 971085, A	c 743	65.2	7.4	5739	48	US-09-619-643-3569	Sequence 3569, Ap
c 671	65.6	7.5	980	46	US-10-301-480C-971085	Sequence 971085, A	c 744	65.2	7.4	5739	48	US-09-619-643-3569	Sequence 3569, Ap
c 672	65.6	7.5	2300	86	US-60-762-056-26311	Sequence 26311, A	c 745	65.2	7.4	5739	48	US-11-227-183A-3569	Sequence 3569, Ap
c 673	65.6	7.5	8467	48	US-10-311-455-82	Sequence 82, Appl	c 746	65.2	7.4	5739	48	US-11-227-183A-3569	Sequence 3569, Ap
c 674	65.6	7.5	8991	42	US-10-221-613-256	Sequence 256, App	c 747	65.2	7.4	5739	48	US-09-666-355A-1335	Sequence 1335, Ap
c 675	65.6	7.5	73370	62	US-10-918-754-16482	Sequence 16482, A	c 748	65.2	7.4	5739	48	US-11-331-032-1335	Sequence 1335, Ap
c 676	65.6	7.5	73370	83	US-60-495-114-16482	Sequence 16482, A	c 749	65.2	7.4	5739	48	US-11-331-032-1335	Sequence 1335, Ap
c 677	65.4	7.5	466	26	US-09-565-306-29077	Sequence 29077, A	c 750	65.2	7.4	5739	48	US-60-207-458-47109	Sequence 47109, A

751	64.8	7.4	800	79	US-60-211-750-74192	Sequence 74192, A	824	64.6	7.4	4604723	36	US-09-947-916-240	Sequence 240, App
C 752	64.8	7.4	890	84	US-60-579-062-33871	Sequence 33871, A	825	64.4	7.4	521	28	US-09-669-817A-17370	Sequence 17370, A
C 753	64.8	7.4	890	85	US-60-603-421-33871	Sequence 33871, A	826	64.4	7.4	649	23	US-09-474-435A-63052	Sequence 63052, A
C 754	64.8	7.4	1076	42	US-10-266-090-33465	Sequence 33465, A	C 827	64.4	7.4	911	84	US-60-579-062-15911	Sequence 15911, A
C 755	64.8	7.4	1099	42	US-10-266-090-33465	Sequence 33465, A	C 828	64.4	7.4	911	85	US-60-603-421-15911	Sequence 15911, A
C 756	64.8	7.4	1114	42	US-10-266-090-19562	Sequence 19562, A	C 829	64.4	7.4	923	79	US-60-207-458-34124	Sequence 34124, A
C 757	64.8	7.4	1153	42	US-10-266-090-27006	Sequence 27006, A	C 830	64.4	7.4	1067	42	US-10-266-090-22100	Sequence 22100, A
C 758	64.8	7.4	1176	39	US-10-015-127-2386	Sequence 2386, A	C 831	64.4	7.4	1130	50	US-10-425-115-132399	Sequence 132399, A
C 759	64.8	7.4	2300	86	US-60-762-056-9392	Sequence 9392, App	C 832	64.4	7.4	1352	42	US-10-266-090-36461	Sequence 36461, A
C 760	64.8	7.4	5464	42	US-10-240-454-38	Sequence 38, App	C 833	64.4	7.4	8310	28	US-09-673-333-9	Sequence 9, Appli
C 761	64.8	7.4	5464	42	US-10-240-454-38	Sequence 38, App	C 834	64.4	7.4	8310	61	US-10-861-875-9	Sequence 9, Appli
C 762	64.8	7.4	5467	48	US-10-311-455-1571	Sequence 1571, App	C 835	64.4	7.4	13511	48	US-10-311-455-254	Sequence 254, App
C 763	64.8	7.4	6381	48	US-10-311-455-1114	Sequence 1114, App	C 836	64.4	7.4	18154	48	US-10-311-455-228	Sequence 228, App
C 764	64.8	7.4	10034	48	US-10-311-455-1857	Sequence 1857, App	C 837	64.4	7.4	50000	54	US-10-706-635-23	Sequence 23, Appl
C 765	64.8	7.4	15282	42	US-10-221-613-97	Sequence 97, Appl	C 838	64.4	7.4	192014	80	US-60-248-823-33	Sequence 33, Appl
C 766	64.6	7.4	536	36	US-09-948-947-1193	Sequence 1193, App	C 839	64.2	7.3	381	28	US-09-654-617-181321	Sequence 181321, A
C 767	64.6	7.4	536	61	US-10-893-315-1193	Sequence 1193, App	C 840	64.2	7.3	381	28	US-09-684-016-181321	Sequence 181321, A
C 768	64.6	7.4	674	22	US-09-304-517A-223802	Sequence 223802, A	C 841	64.2	7.3	548	28	US-09-654-617-83820	Sequence 83820, A
C 769	64.6	7.4	674	22	US-09-371-146A-223802	Sequence 223802, A	C 842	64.2	7.3	548	28	US-09-684-016-83820	Sequence 83820, A
C 770	64.6	7.4	674	38	US-09-985-678-223802	Sequence 223802, A	C 843	64.2	7.3	612	65	US-11-031-175-1357	Sequence 1357, App
C 771	64.6	7.4	674	78	US-60-125-816-7377	Sequence 7377, App	C 844	64.2	7.3	711	28	US-09-666-355A-1218	Sequence 1218, App
C 772	64.6	7.4	679	23	US-09-474-435A-43816	Sequence 43816, A	C 845	64.2	7.3	711	73	US-11-331-032-1218	Sequence 1218, App
C 773	64.6	7.4	898	42	US-10-266-090-33207	Sequence 33207, A	C 846	64.2	7.3	1127	42	US-10-266-090-17252	Sequence 17252, A
C 774	64.6	7.4	1010	33	US-09-881-797-607	Sequence 607, App	C 847	64.2	7.3	1158	42	US-10-266-090-18343	Sequence 18343, A
C 775	64.6	7.4	1010	39	US-10-021-698-607	Sequence 607, App	C 848	64.2	7.3	1255	42	US-10-266-090-22716	Sequence 22716, A
C 776	64.6	7.4	1242	74	US-11-360-355-32043	Sequence 32043, A	C 849	64.2	7.3	6292	42	US-10-221-714A-462	Sequence 462, App
C 777	64.6	7.4	1242	85	US-60-655-875-32043	Sequence 32043, A	C 850	64.2	7.3	6292	42	US-10-221-714A-462	Sequence 462, App
C 778	64.6	7.4	13859	42	US-10-266-090-24327	Sequence 24327, A	C 851	64.2	7.3	18218	48	US-10-311-455-1921	Sequence 1921, App
C 779	64.6	7.4	13859	50	US-10-498-430-125	Sequence 125, App	C 852	64.2	7.3	597573	36	US-09-947-911-179	Sequence 179, App
C 780	64.6	7.4	5286	50	US-10-498-430-241	Sequence 241, App	C 853	64	7.3	467	28	US-09-654-617-181119	Sequence 181119, A
C 781	64.6	7.4	5286	51	US-10-517-441-294	Sequence 294, App	C 854	64	7.3	467	28	US-09-684-016-181119	Sequence 181119, A
C 782	64.6	7.4	5286	51	US-10-517-441-294	Sequence 294, App	C 855	64	7.3	470	33	US-09-873-402A-1978	Sequence 1978, App
C 783	64.6	7.4	5945	48	US-10-311-455-57	Sequence 57, App	C 856	64	7.3	470	79	US-60-209-830-1978	Sequence 1978, App
C 784	64.6	7.4	7001	41	US-10-172-086-59	Sequence 59, Appl	C 857	64	7.3	489	28	US-09-669-817A-32071	Sequence 32071, A
C 785	64.6	7.4	7001	41	US-10-172-086A-59	Sequence 59, Appl	C 858	64	7.3	547	33	US-09-873-402A-1402	Sequence 1402, App
C 786	64.6	7.4	7001	41	US-10-172-086A-59	Sequence 59, Appl	C 859	64	7.3	547	79	US-60-209-830-1402	Sequence 1402, App
C 787	64.6	7.4	7001	48	US-10-311-506A-1	Sequence 1, Appli	C 860	64	7.3	1232	42	US-10-266-090-32307	Sequence 32307, A
C 788	64.6	7.4	7001	48	US-10-311-506A-1	Sequence 1, Appli	C 861	64	7.3	1240	42	US-10-266-090-35800	Sequence 35800, A
C 789	64.6	7.4	7001	48	US-10-311-507-1	Sequence 1, Appli	C 862	64	7.3	7676	42	US-10-240-485-152	Sequence 152, App
C 790	64.6	7.4	7001	50	US-10-480-846-59	Sequence 59, Appl	C 863	64	7.3	32392	54	US-10-706-635-27	Sequence 27, Appl
C 791	64.6	7.4	13859	36	US-09-950-083B-10355	Sequence 10355, A	C 864	64	7.3	32392	69	US-11-121-086-25	Sequence 25, Appl
C 792	64.6	7.4	13859	36	US-09-950-083B-10355	Sequence 10355, A	C 865	64	7.3	224986	80	US-60-261-974-42	Sequence 42, Appl
C 793	64.6	7.4	13859	41	US-10-105-299-12506	Sequence 12506, A	C 866	64	7.3	335913	30	US-09-754-853A-2	Sequence 2, Appli
C 794	64.6	7.4	13859	61	US-10-868-184A-10355	Sequence 10355, A	C 867	64	7.3	335913	30	US-09-754-853A-3	Sequence 3, Appli
C 795	64.6	7.4	13859	61	US-10-868-184B-10355	Sequence 10355, A	C 868	64	7.3	335913	32	US-09-855-768-5	Sequence 5, Appli
C 796	64.6	7.4	13859	61	US-10-868-184C-10355	Sequence 10355, A	C 869	64	7.3	335913	74	US-11-389-017-2	Sequence 2, Appli
C 797	64.6	7.4	13859	61	US-10-868-184C-10355	Sequence 10355, A	C 870	64	7.3	335913	74	US-11-389-017-3	Sequence 3, Appli
C 798	64.6	7.4	13859	61	US-10-868-184D-10355	Sequence 10355, A	C 871	64	7.3	335913	74	US-11-389-343-2	Sequence 2, Appli
C 799	64.6	7.4	13859	61	US-10-868-184B-10355	Sequence 10355, A	C 872	64	7.3	335913	74	US-11-389-343-2	Sequence 3, Appli
C 800	64.6	7.4	13859	61	US-10-896-164-10355	Sequence 10355, A	C 873	64	7.3	416223	84	US-60-568-845-2888	Sequence 2888, App
C 801	64.6	7.4	18154	48	US-10-240-452-34	Sequence 237, App	C 874	63.8	7.3	419	23	US-09-480-902-11234	Sequence 11234, A
C 802	64.6	7.4	18154	48	US-10-240-452-34	Sequence 34, Appl	C 875	63.8	7.3	419	37	US-09-960-352-11234	Sequence 11234, A
C 803	64.6	7.4	18683	48	US-10-311-455-286	Sequence 286, App	C 876	63.8	7.3	425	32	US-09-837-604A-77902	Sequence 77902, A
C 804	64.6	7.4	47573	78	US-60-119-494-1	Sequence 1, Appli	C 877	63.8	7.3	425	32	US-09-837-604B-77902	Sequence 77902, A
C 805	64.6	7.4	98573	80	US-60-218-505-542	Sequence 542, App	C 878	63.8	7.3	425	78	US-60-197-872-73058	Sequence 73058, A
C 806	64.6	7.4	134987	36	US-09-949-003C-4984	Sequence 4984, App	C 879	63.8	7.3	452	33	US-09-873-402A-89832	Sequence 89832, A
C 807	64.6	7.4	134987	36	US-09-949-003C-5015	Sequence 5015, App	C 880	63.8	7.3	452	79	US-60-209-830-61571	Sequence 61571, A
C 808	64.6	7.4	134987	62	US-10-940-774-15348	Sequence 15348, A	C 881	63.8	7.3	483	85	US-60-680-473-1911	Sequence 1911, App
C 809	64.6	7.4	134987	62	US-10-940-774-15348	Sequence 15349, A	C 882	63.8	7.3	483	85	US-60-544-1911	Sequence 1911, App
C 810	64.6	7.4	134987	62	US-10-940-774-15350	Sequence 15350, A	C 883	63.8	7.3	716	84	US-60-579-062-25379	Sequence 25379, A
C 811	64.6	7.4	134987	62	US-10-940-774-15507	Sequence 15507, A	C 884	63.8	7.3	716	85	US-60-603-421-25379	Sequence 25379, A
C 812	64.6	7.4	134987	62	US-10-940-774-15508	Sequence 15508, A	C 885	63.8	7.3	723	26	US-09-565-308-26980	Sequence 26980, A
C 813	64.6	7.4	134987	62	US-10-940-774-15509	Sequence 15509, A	C 886	63.8	7.3	723	74	US-11-353-150-26980	Sequence 26980, A
C 814	64.6	7.4	134987	63	US-10-940-774A-15348	Sequence 15348, A	C 887	63.8	7.3	812	42	US-10-266-090-25616	Sequence 25616, A
C 815	64.6	7.4	134987	63	US-10-940-774A-15349	Sequence 15349, A	C 888	63.8	7.3	833	84	US-60-579-062-21616	Sequence 21616, A
C 816	64.6	7.4	134987	63	US-10-940-774A-15350	Sequence 15350, A	C 889	63.8	7.3	833	85	US-60-603-421-21616	Sequence 21616, A
C 817	64.6	7.4	134987	63	US-10-940-774A-15507	Sequence 15507, A	C 890	63.8	7.3	878	84	US-60-579-062-35648	Sequence 35648, A
C 818	64.6	7.4	134987	63	US-10-940-774A-15508	Sequence 15508, A	C 891	63.8	7.3	878	85	US-60-603-421-35648	Sequence 35648, A
C 819	64.6	7.4	134987	63	US-10-940-774A-15509	Sequence 15509, A	C 892	63.8	7.3	990	84	US-60-579-062-19101	Sequence 19101, A
C 820	64.6	7.4	142987	82	US-60-465-241-52528	Sequence 52528, A	C 893	63.8	7.3	990	85	US-60-603-421-19101	Sequence 19101, A
C 821	64.6	7.4	156634	64	US-10-930-328-94060	Sequence 94060, A	C 894	63.8	7.3	1008	28	US-09-666-355A-1062	Sequence 1062, App
C 822	64.6	7.4	156634	64	US-10-930-328A-94060	Sequence 94060, A	C 895	63.8	7.3	1008	73	US-11-331-032-1062	Sequence 1062, App
C 823	64.6	7.4	3710966	36	US-09-948-128-368	Sequence 368, App	C 896	63.8	7.3	1125	42	US-10-266-090-32115	Sequence 32115, A





```

Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT 120
QY 121 CTTGCTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACAT 180
Db 121 CTTGCTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACAT 180
QY 181 ACTTTAAACATTTAATCCAAAAATCAAGAAATGAGACCGAATCAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAAATCAAGAAATGAGACCGAATCAATTTATGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAACTTTTACGATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAACTTTTACGATATTTTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTTAATGCTTTATTTAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTTAATGCTTTATTTAA 360
QY 361 TTTCAACGAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCAACGAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
QY 541 GATGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Db 541 GATGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAATCTTACAAAAATAAATCTATATGCTCAATCAAGAAATTTAATGCTTTTAA 660
Db 601 GAAAAAATCTTACAAAAATAAATCTATATGCTCAATCAAGAAATTTAATGCTTTTAA 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATACTACACT 720
QY 721 AAAGATATACCTCATACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTAA 780
Db 721 AAAGATATACCTCATACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTAA 780

```

RESULT 2

```

US-10-734-719-8
; Sequence 8, Application US/10734719
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/734,719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21

```

```

; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:19
US-10-734-719-8

```

```

Query Match 100.0%; Score 876; DB 60; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAGATAAATACAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAGATAAATACAT 120
QY 121 CTTGCTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACAT 180
Db 121 CTTGCTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACAT 180
QY 181 ACTTTAAACATTTAATCCAAAAATCAAGAAATGAGACCGAATCAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAAATCAAGAAATGAGACCGAATCAATTTATGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAACTTTTACGATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAACTTTTACGATATTTTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTTAATGCTTTTAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTTAATGCTTTTAA 360
QY 361 TTTCAACGAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCAACGAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
QY 541 GATGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Db 541 GATGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAATCTTACAAAAATAAATCTATATGCTCAATCAAGAAATTTAATGCTTTTAA 660
Db 601 GAAAAAATCTTACAAAAATAAATCTATATGCTCAATCAAGAAATTTAATGCTTTTAA 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATACTACACT 720
QY 721 AAAGATATACCTCATACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTAA 780
Db 721 AAAGATATACCTCATACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTAA 780

```



```
Qy 781 AAAAAAAAAAATAAGCAATTTATTTCAAAGGAAATAA 876
Db 781 AAAAAAAAAAATAAGCAATTTATTTCAAAGGAAATAA 840
Qy 841 CCTAGTATATAAGCAATTTATTTCAAAGGAAATAA 876
Db 841 CCTAGTATATAAGCAATTTATTTCAAAGGAAATAA 876

RESULT 3
US-10-820-536-8
; Sequence 8, Application US/10820536
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:19
US-10-820-536-8

Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTTATTGCTGGAATCGCAAGTTTAAAAAGAAATTTGATTATTAAGATTA 840
Db 1 ATGAAAAAGCTTATTGCTGGAATCGCAAGTTTAAAAAGAAATTTGATTATTAAGATTA 840
Qy 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAAAATTAATCCAAAATCAAGAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAATTAATCCAAAATCAAGAATATGAGACCGAACTAATTTATGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCT 300
Qy 301 GATGCTCAATTTGGGATATGATTTTAAAAAATTTTAAAAAATTTTAAAGAAATTTAAAT 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAAAAATTTTAAAAAATTTTAAAGAAATTTAAAT 360
Qy 361 TTTTCAGAAATTTATTTCAATCAAGAATTAACCTCAGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAGAATTAACCTCAGGGTCTATATGTTGCGAGTAGCC 420
```

```
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAACTTACAAAAATAAACTATTATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAACTTACAAAAATAAACTATTATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 660
Qy 661 ATAGAACCTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACCT 720
Db 661 ATAGAACCTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACCT 720
Qy 721 AAAGATATATCATACCTTCTAGTGAGCTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATATCATACCTTCTAGTGAGCTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAAATAAAAAATTTAAAGAAATTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTTAAAGAAATTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTATATAAGCAATTTATTTCAAAGGAAATAA 876
Db 841 CCTAGTATATAAGCAATTTATTTCAAAGGAAATAA 876
```

## RESULT 4

```
US-10-821-573-8
; Sequence 8, Application US/10821573
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:19
US-10-821-573-8
```

```
Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTTATTGCTGGAATCGCAAGTTTAAAAAGAAATTTGATTATTAAGATTA 60
Db 1 ATGAAAAAGCTTATTGCTGGAATCGCAAGTTTAAAAAGAAATTTGATTATTAAGATTA 60
```

```
Db 1 ATGAAAAAGTTATTATTGCTGGAATGSAACCAAGTTTTAAAGAAATGATTATTCAAGG 60
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAAATCAAGAAATAGAGACCGAACTAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATAGAGACCGAACTAATATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTATGATACCAAGAAATTTCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTATGATACCAAGAAATTTCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTTCAAAATAAAACTATATGTTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTTCAAAATAAAACTATATGTTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACT 720
Qy 721 AAAGATATACTATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATACTATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTAATTTT 780
Qy 781 AAAAAATTAATAAATTTAAGAAAAATGTTTATTAAGATTTGATAAAGATCTATTAAGATTA 840
Db 781 AAAAAATTAATAAATTTAAGAAAAATGTTTATTAAGATTTGATAAAGATCTATTAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAAATTA 876
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAAATTA 876
```

RESULT 5

US-10-821-604-8  
; Sequence 8, Application US/10821604  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; PRIOR FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstrII) from C. jejuni O:19  
US-10-821-604-8

```
Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGSAACCAAGTTTTAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGSAACCAAGTTTTAAAGAAATGATTATTCAAGG 60
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAAATCAAGAAATAGAGACCGAACTAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATAGAGACCGAACTAATATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTATGATACCAAGAAATTTCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTATGATACCAAGAAATTTCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTTCAAAATAAAACTATATGTTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTTCAAAATAAAACTATATGTTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACT 720
Qy 721 AAAGATATACTATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATACTATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTAATTTT 780
```

Db 721 AAAGATATCTACTATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
Qy 781 AAAAAATAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTAATTTCAAAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCAATTAATTTCAAAAGGAAAAATAA 876

RESULT 6  
US-10-830-825-8  
; Sequence 8, Application US/10830825  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/830,825  
; CURRENT FILING DATE: 2004-04-24  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-830-825-8

Query Match 100.0%; Score 876; DB 61; Length 876;  
Best Local Similarity 100.0%; Pred. No. 3.4e-192;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60  
Db 1 ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGACCAATAC 180  
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGACCAATAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGCACTAATATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGCACTAATATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTTCAGATATTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTTCAGATATTTTCCT 300

Qy 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360

Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420  
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGTATACCAAAACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAAAACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

Qy 601 GAAAAAATCTTACAAAAATAAACTATATGCTTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATCTTACAAAAATAAACTATATGCTTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660

Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720

Qy 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780

Qy 781 AAAAAATAAAATTTAAAAAGAAATCTTTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAAATTTAAAAAGAAATCTTTTATTACAAAGTTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAAATAA 876

## RESULT 7

US-10-830-997-8  
; Sequence 8, Application US/10830997  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/830,997  
; CURRENT FILING DATE: 2004-04-22  
; PRIOR APPLICATION NUMBER: US/10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-830-997-8

Query Match 100.0%; Score 876; DB 61; Length 876;  
Best Local Similarity 100.0%; Pred. No. 3.4e-192;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60  
Db 1 ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60

```
Db 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGTTTTAAAGAAATGATTATTCAAGG 60
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTAAT 240
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTABAACTTTTATGAGATTTTTCCT 300
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTABAACTTTTATGAGATTTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720
Qy 721 AAAGATATACATACTACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATACATACTACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
Qy 781 AAAAAATTAATAAATAAAGAAATTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAAATAAAGAAATTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTATATAAGCATTTTCAAGGAAAAATA 876
Db 841 CCTAGTATATAAGCATTTTCAAGGAAAAATA 876
```

## RESULT 8

```
US-10-845-408-8
; Sequence 8, Application US/10845408
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
```

```
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-845-408-8
```

```
Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Qy 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATACTAC 180
Qy 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTAAT 240
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTABAACTTTTACGATTTTTCCT 300
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTABAACTTTTACGATTTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720
Qy 721 AAAGATATACATACTACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATACATACTACCTTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
```

```
Qy 781 AAAAAATAAAAAATTAAGAAATGTTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAAATTAAGAAATGTTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGAGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGAGAAATAA 876

RESULT 9
US-10-845-412-8
; Sequence 8, Application US/10845412
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-845-412-8

Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTCTTTTGACCACTACTAC 180
Db 121 CTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTCTTTTGACCACTACTAC 180
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGCACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGCACTAATTTATGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360
Qy 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGCTGCGAGTAGCC 420
Db 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGCTGCGAGTAGCC 420
```

```
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTGCACTATATCGGACATAGTAAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTGCACTATATCGGACATAGTAAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTACCAAAATTTT 660
Db 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTACCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 720
Qy 721 AAAGATATATCATACACTTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATATCATACACTTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTAATTTT 780
Qy 781 AAAAAAATAAAAAATTAAGAAATGCTTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTAAGAAATGCTTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGAGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGAGAAATAA 876
```

## RESULT 10

```
US-10-846-219-8
; Sequence 8, Application US/10846219
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-846-219-8
```

```
Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
Db 1 ATGAAAAAGTTATTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
```

```
Qy 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATAC 180
Qy 181 ACTTTAAACATTTAAATCCRAAATCAAGAATATGAGCCGACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAAATCCRAAATCAAGAATATGAGCCGACTAATTTATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGAAGAACTTTTACGATTTATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGAAGAACTTTTACGATTTATTTTCT 300
Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTCAAGAAATTTTCAATCAAGAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCAAGAAATTTTCAATCAAGAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGTCA 480
Qy 481 TCTTATGCTTTTGATACAAAGAAATTTTAAACTAGCCCTGATTTTAAAGAT 540
Db 481 TCTTATGCTTTTGATACAAAGAAATTTTAAACTAGCCCTGATTTTAAAGAT 540
Qy 541 GATCGCTGCTATATGCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTGCTATATGCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTCAAAATTAATACTATTTGCTTATGCTCTTAAGTCTTTTACCAATTTT 660
Db 601 GAAAAAATTTCAAAATTAATACTATTTGCTTATGCTCTTAAGTCTTTTACCAATTTT 660
Qy 661 ATAGAACTTAGCGCCAAATTTAAATTTCAATTTTATCACAAGAAAAATACTACACT 720
Db 661 ATAGAACTTAGCGCCAAATTTAAATTTCAATTTTATCACAAGAAAAATACTACACT 720
Qy 721 AAAGATATCTATACCTCTTAGTGAGCTTTATGAAAATTTTCAAAAATTTAATTTT 780
Db 721 AAAGATATCTATACCTCTTAGTGAGCTTTATGAAAATTTTCAAAAATTTAATTTT 780
Qy 781 AAAAAATTAATAATTAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
```

RESULT 11

```
US-10-847-983-8
; Sequence 8, Application US/10847983
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
```

```
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-847-983-8
```

```
Query Match 100.0%; Score 876; DB 61; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTATTCTCGAAATGACCAAGTTTAAAGAAAAATTTGATTATTTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCTCGAAATGACCAAGTTTAAAGAAAAATTTGATTATTTCAAGG 60
Qy 61 CTACCAATGATTTTGGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATAC 180
Qy 181 ACTTTAAACATTTAAATCCRAAATCAAGAATATGAGCCGACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAAATCCRAAATCAAGAATATGAGCCGACTAATTTATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTCAAGAAATTTTCAATCAAGAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCAAGAAATTTTCAATCAAGAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTGCTATATGCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTGCTATATGCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTCAAAATTAATACTATTTGCTTATGCTCTTAAGTCTTTTACCAATTTT 660
Db 601 GAAAAAATTTCAAAATTAATACTATTTGCTTATGCTCTTAAGTCTTTTACCAATTTT 660
Qy 661 ATAGAACTTAGCGCCAAATTTAAATTTCAATTTTATCACAAGAAAAATACTACACT 720
Db 661 ATAGAACTTAGCGCCAAATTTAAATTTCAATTTTATCACAAGAAAAATACTACACT 720
Qy 721 AAAGATATCTATACCTCTTAGTGAGCTTTATGAAAATTTTCAAAAATTTAATTTT 780
Db 721 AAAGATATCTATACCTCTTAGTGAGCTTTATGAAAATTTTCAAAAATTTAATTTT 780
Qy 781 AAAAAATTAATAATTAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
```

Db 781 AAAAAAAAAAATAAAGAAATGTTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAATAA 876  
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAATAA 876

RESULT 12  
US-10-850-125-8  
; Sequence 8, Application US/10850125  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,125  
; CURRENT FILING DATE: 2004-05-19  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19

US-10-850-125-8  
Query Match 100.0%; Score 876; DB 61; Length 876;  
Best Local Similarity 100.0%; Pred. No. 3.4e-192;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATCGCAAGTTTAAAGAAATGATTTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATCGCAAGTTTAAAGAAATGATTTTCAAGG 60

Qy 61 CTACCAATCATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180  
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180

Qy 181 ACTTTAAAAATTAATTCRAAATCAAGATATGAGACCGAATTAATGTTGTTCAAT 240  
Db 181 ACTTTAAAAATTAATTCRAAATCAAGATATGAGACCGAATTAATGTTGTTCAAT 240

Qy 241 TACACCAAGCTCATCTAGAAAAAGAAATTTTGAAAACTTTTACGATTATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAAAGAAATTTTGAAAACTTTTACGATTATTTTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTTTAAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360

Qy 361 TTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGAGTAGCC 420  
Db 361 TTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGAGTAGCC 420

Qy 421 ATAGCCCTAGGATCAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATCAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAAGCTTTAGAAATTTCTA 600

Qy 601 GAAAAAACTTACAAAAATAAACTATATTGCTTATGCTTATGCTTCTTAATAGTCTTTTACCAATTTT 660  
Db 601 GAAAAAACTTACAAAAATAAACTATATTGCTTATGCTTATGCTTCTTAATAGTCTTTTACCAATTTT 660

Qy 661 ATAGAAGCTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATACACACT 720  
Db 661 ATAGAAGCTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATACACACT 720

Qy 721 AAAGATATACATACACTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780  
Db 721 AAAGATATACATACACTCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780

Qy 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876  
Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876

RESULT 13  
US-10-850-807-8  
; Sequence 8, Application US/10850807  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,807  
; CURRENT FILING DATE: 2004-05-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19

US-10-850-807-8  
Query Match 100.0%; Score 876; DB 61; Length 876;  
Best Local Similarity 100.0%; Pred. No. 3.4e-192;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATCGCAAGTTTAAAGAAATGATTTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATCGCAAGTTTAAAGAAATGATTTTCAAGG 60

Qy 61 CTACCAATGATTTTGATGATTTTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGATGATTTTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

```
QY 121 CTTGGTAAAAATGCAAGAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACCTAC 180
Db 121 CTTGGTAAAAATGCAAGAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACCTAC 180
QY 181 ACTTTAAACAATTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTTAAT 240
Db 181 ACTTTAAACAATTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
QY 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTGCTAGTACC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTGCTAGTACC 420
QY 421 ATAGCCCTAGGATCAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATCAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Db 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
QY 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTTAAATTT 780
Db 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTTAAATTT 780
QY 781 AAAAAATTTAAATTTAAAGAAATTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTTAAATTTAAAGAAATTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
```

## RESULT 14

US-10-961-882-8

```
; Sequence 8, Application US/10961882
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/961,882
; PRIOR FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-961-882-8

Query Match      100.0%; Score 876; DB 63; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGATTATTTCTGCGAATGACCAAGTTTAAAGAAATTTAAAGAAATTTAAAGG 60
Db 1 ATGAAAAAGATTATTTCTGCGAATGACCAAGTTTAAAGAAATTTAAAGAAATTTAAAGG 60
QY 61 CTACCAAAATGATTTTATGATGATTTAGATGATTAATCAATTTTATTTGAAAGATAAATAC 120
Db 61 CTACCAAAATGATTTTATGATGATTTAGATGATTAATCAATTTTATTTGAAAGATAAATAC 120
QY 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACGCCCTAAATTTCTTTGAGCAATACCTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACGCCCTAAATTTCTTTGAGCAATACCTAC 180
QY 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTAAA 360
QY 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTGAGTAGGCC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTGAGTAGGCC 420
QY 421 ATAGCCCTAGGATCAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATCAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAAAT 540
QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Db 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
QY 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTTAAATTT 780
Db 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTTAAATTT 780
QY 781 AAAAAATTTAAATTTAAAGAAATTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTTAAATTTAAAGAAATTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
```



```
Db      841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876

RESULT 15
US-10-962-235-8
; Sequence 8, Application US/10962235
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/962,235
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:19
US-10-962-235-8

Query Match      100.0%; Score 876; DB 63; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60
Db      1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60

Qy      61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAGATAAATACTAT 120
Db      61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAGATAAATACTAT 120

Qy      121 CTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db      121 CTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180

Qy      181 ACTTTAAAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
Db      181 ACTTTAAAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240

Qy      241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300
Db      241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300

Qy      301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
Db      301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360

Qy      361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db      361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy      421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db      421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480

Qy      481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db      481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

Qy      541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTTA 600
Db      541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTTA 600

Qy      601 GAAAAAACTTTACAAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 660
Db      601 GAAAAAACTTTACAAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 660

Qy      661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720
Db      661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720

Qy      721 AAAGATATACTCATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
Db      721 AAAGATATACTCATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780

Qy      781 AAAAAAATAAAAAATTAAGAAAAATTTTATTAACAAGTTGATAAAAAGATCTATTAAAGATTA 840
Db      781 AAAAAAATAAAAAATTAAGAAAAATTTTATTAACAAGTTGATAAAAAGATCTATTAAAGATTA 840

Qy      841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876
Db      841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876

RESULT 16
US-10-962-334-8
; Sequence 8, Application US/10962334
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/962,334
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:19
US-10-962-334-8

Query Match      100.0%; Score 876; DB 63; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.4e-192;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60
Db      1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60

Qy      61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAGATAAATACTAT 120
Db      61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAGATAAATACTAT 120

Qy      121 CTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db      121 CTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180

Qy      181 ACTTTAAAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
Db      181 ACTTTAAAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240

Qy      241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300
Db      241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300

Qy      301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
Db      301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360

Qy      361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db      361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy      421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db      421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480

Qy      481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db      481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
```

```
Qy 181 ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAAATATATGTCTTAAT 240
Db 181 ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAAATATATGTCTTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCCT 300
Qy 301 GATCCTCATTTGGGATATGATTTTTTAAACAATCAAGATATTAAGATTTTAAAT 360
Db 301 GATCCTCATTTGGGATATGATTTTTTAAACAATCAAGATATTAAGATTTTAAAT 360
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTGCGAGTAGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTAAAT 540
Qy 541 GATCGCTCGACTATATCGGACATAGTAAATCAAGATATCAAGATATTAAGATTTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAATCAAGATATCAAGATATTAAGATTTTCTA 600
Qy 601 GAAAAAATTTTCAAAATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 660
Db 601 GAAAAAATTTTCAAAATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATCAAAATTTTATATGCTTCTTAATAGCAATTTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATCAAAATTTTATATGCTTCTTAATAGCAATTTT 720
Qy 721 AAAGATATCTATCATCTCTAGTGGGCTTTATGAGGCTTTTCAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATCATCTCTAGTGGGCTTTATGAGGCTTTTCAAAATTTTAAATTTT 780
Qy 781 AAAAAATTAAGATATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 840
Db 781 AAAAAATTAAGATATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 840
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
```

## RESULT 17

```
US-10-303-161-13
; Sequence 13, Application US/10303161
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CetII) from C. jejuni O:36
US-10-303-161-13
```

```
Query Match 99.3%; Score 869.8; DB 47; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTTATTATTCTGCGAAATGAGCCAAAGTTTAAAGAAATGATTTTCAAG 60
Db 1 ATGAAAAAGTTATTATTCTGCGAAATGAGCCAAAGTTTAAAGAAATGATTTTCAAG 60
Qy 61 CTACCAATGATTTTGTATGATTTTATAGATGATTAATCAATTTTAAAGATTAATTA 120
Db 61 CTACCAATGATTTTGTATGATTTTATAGATGATTAATCAATTTTAAAGATTAATTA 120
Qy 121 CTTCGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAACTAAATATGCTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAACTAAATATGCTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAATTTTAAAGAAATTTTAAAGATTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAATTTTAAAGAAATTTTAAAGATTTTAA 360
Qy 361 TTTACGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGCTGAGTAGCC 420
Db 361 TTTACGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGCTGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
Qy 541 GATCGCTCGACTATATCGGACATAGTAAATCAAGATATCAAGATATTAAGATTTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAATCAAGATATCAAGATATTAAGATTTTCTA 600
Qy 601 GAAAAAATTTTCAAAATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 660
Db 601 GAAAAAATTTTCAAAATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATCAAAATTTTATATGCTTCTTAATAGCAATTTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATCAAAATTTTATATGCTTCTTAATAGCAATTTT 720
Qy 721 AAAGATATCTATCATCTCTAGTGGGCTTTATGAGGCTTTTCAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATCATCTCTAGTGGGCTTTATGAGGCTTTTCAAAATTTTAAATTTT 780
Qy 781 AAAAAATTAAGATATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 840
Db 781 AAAAAATTAAGATATTAAGATATTTTAACTAGTAAATCAAGATATTAAGATTTTCTA 840
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
```

RESULT 18

```
US-10-734-719-13
; Sequence 13, Application US/10734719
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10734, 719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-734-719-13

Query Match      99.3%; Score 869.8; DB 60; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAAATGATTATTCAAGG 60
DB      1  ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAAATGATTATTCAAGG 60

QY      61  CTACCAATGATTTGATGTAATCAATTTTATTTGAAGATAAACTACTAT 120
DB      61  CTACCAATGATTTGATGTAATCAATTTTATTTGAAGATAAACTACTAT 120

QY      121  CTTCGTAATAAATGCAAGACAGTGTTTACACCCCTAATTTCTTCTTGAGCAACTACTAC 180
DB      121  CTTCGTAATAAATGCAAGACAGTGTTTACACCCCTAATTTCTTCTTGAGCAACTACTAC 180

QY      181  ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACCTAATTTATGTGTTCTAAT 240
DB      181  ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACCTAATTTATGTGTTCTAAT 240

QY      241  TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
DB      241  TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300

QY      301  GATGCTCAATTTGGGATGATGTTTTTAAACCACTTAAAGAAATTAATGCTTTATTTAAA 360
DB      301  GATGCTCAATTTGGGATGATGTTTTTAAACCACTTAAAGAAATTAATGCTTTATTTAAA 360

QY      361  TTTTCACCAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC 420
DB      361  TTTTCACCAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC 420

QY      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY      481  TCTTATGCTTTTGATACAAACAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540
DB      481  TCTTATGCTTTTGATACAAACAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540

QY      541  GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB      541  GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

US-10-820-536-13
; Sequence 13, Application US/10820536
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-820-536-13

Query Match      99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAAATGATTATTCAAGG 60
DB      1  ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAAATGATTATTCAAGG 60

QY      61  CTACCAATGATTTGATGTAATCAATTTTATTTGAAGATAAACTACTAT 120
DB      61  CTACCAATGATTTGATGTAATCAATTTTATTTGAAGATAAACTACTAT 120

QY      121  CTTCGTAATAAATGCAAGACAGTGTTTACACCCCTAATTTCTTCTTGAGCAACTACTAC 180
DB      121  CTTCGTAATAAATGCAAGACAGTGTTTACACCCCTAATTTCTTCTTGAGCAACTACTAC 180

QY      181  ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACCTAATTTATGTGTTCTAAT 240
DB      181  ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACCTAATTTATGTGTTCTAAT 240

QY      241  TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
DB      241  TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300

QY      301  GATGCTCAATTTGGGATGATGTTTTTAAACCACTTAAAGAAATTAATGCTTTATTTAAA 360
DB      301  GATGCTCAATTTGGGATGATGTTTTTAAACCACTTAAAGAAATTAATGCTTTATTTAAA 360

QY      361  TTTTCACCAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC 420
DB      361  TTTTCACCAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC 420

QY      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY      481  TCTTATGCTTTTGATACAAACAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540
DB      481  TCTTATGCTTTTGATACAAACAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540

QY      541  GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB      541  GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
```

```
Db 181 ACTTTAAAAACATTAATCCAAATACAGATATGAGACCGAACTAATATATGTGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTGTAAGAACTTTTAAAGAACTTTTACGATATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTGTAAGAACTTTTAAAGAACTTTTACGATATTTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTTAAA 360
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAGAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAGAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACTATATCGGATAGTAAAGAAATACAGATATAAGAAAGCTTTTAAAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGATAGTAAAGAAATACAGATATAAGAAAGCTTTTAAAGAAATTTCTA 600
Qy 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660
Db 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAG 720
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAG 720
Qy 721 AAAGATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 780
Db 721 AAAGATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 780
Qy 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840
Db 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
```

RESULT 20

```
US-10-821-573-13
; Sequence 13, Application US/10821573
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; PRIORITY FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-821-573-13

Query Match 99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTTATTATTCCTCGAAATGGAACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60
Db 1 ATGAAAAAGTTTATTATTCCTCGAAATGGAACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60
Qy 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTTAAT 240
Qy 241 TACAAACAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Db 241 TACAAACAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAAAT 540
Qy 541 GATCGCTCGACTATATCGGATAGTAAAGAAATACAGATATAAGAAAGCTTTTAAAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGATAGTAAAGAAATACAGATATAAGAAAGCTTTTAAAGAAATTTCTA 600
Qy 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660
Db 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAG 720
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAG 720
Qy 721 AAAGATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 780
Db 721 AAAGATATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 780
Qy 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840
Db 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
```

```
US-10-821-604-13
; Sequence 13, Application US/10821604
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,604
; CURRENT FILING DATE: 2004-04-08
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-821-604-13

Query Match          99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60
Db 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60

Qy 61 CTACCAAAATGATTTGATGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTGATGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 121 CTGTGTAATAAATGCAAGCGAGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGCGAGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACCTAATTAATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACCTAATTAATGTGTTCTAAT 240

US-10-821-604-13
; Sequence 13, Application US/10830825
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/830,825
; CURRENT FILING DATE: 2004-04-24
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-830-825-13

Query Match          99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60
Db 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60

Qy 61 CTACCAAAATGATTTGATGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTGATGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 121 CTGTGTAATAAATGCAAGCGAGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGCGAGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACCTAATTAATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACCTAATTAATGTGTTCTAAT 240
```



```
US-10-845-408-13
; Sequence 13, Application US/10845408
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-408-13
Query Match          99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTATTTGAAGATAAATCTAT 120
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTATTTGAAGATAAATCTAT 120

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120

Qy 121 CTGTTGTAATAATGCAAGACAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATCTAT 180
Db 121 CTGTTGTAATAATGCAAGACAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATCTAT 180

Qy 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy 361 TTTTCAACAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTTCAACAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540

Qy 541 GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600

US-10-845-412-13
; Sequence 13, Application US/10845412
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-412-13
Query Match          99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTATTTGAAGATAAATCTAT 120
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATTTATTTGAAGATAAATCTAT 120

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120

Qy 121 CTGTTGTAATAATGCAAGACAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATCTAT 180
Db 121 CTGTTGTAATAATGCAAGACAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATCTAT 180

Qy 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy 361 TTTTCAACAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTTCAACAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

Qy 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540

Qy 541 GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
```

```
Db 181 ACTTTAAACCAATTAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATATTTTCCCT 300
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATATTTTCCCT 300
Qy 301 GATCCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTTAAATGCTTATTTAAA 360
Db 301 GATCCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTTAAATGCTTATTTAAA 360
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Db 601 GAAAAAATTTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720
Qy 721 AAAGATATCTATACACTTTCTAGTGGGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATACACTTTCTAGTGGGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAATTTAAATTTAAAGAAATGTTTATACAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTTAAATTTAAAGAAATGTTTATACAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
```

## RESULT 26

```
US-10-846-219-13
; Sequence 13, Application US/10846219
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-846-219-13
```

```
Query Match 99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTATTATTGCTCGAAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTATTATTGCTCGAAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Qy 61 CTACCAAAATGATTTTGTGCTATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGTGCTATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGTAAAAAATGCAAGACGTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180
Db 121 CTTGTAAAAAATGCAAGACGTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGACCGCACTAATTTATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGACCGCACTAATTTATGTTCTTAAT 240
Qy 241 TACACCAAGCTCATCTGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300
Db 241 TACACCAAGCTCATCTGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300
Qy 301 GATGCTCATTTTGGGATATGATTTTTTAAACCACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTTGGGATATGATTTTTTAAACCACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 361 TTTTCGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Db 361 TTTTCGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTTGAATTTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTTGAATTTTCTA 600
Qy 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Db 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720
Qy 721 AAAGATATCTATACACTTTCTAGTGGGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATACACTTTCTAGTGGGCTTTATGAAAAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAATTTAAATTTAAAGAAATGTTTATACAGTTGATAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTTAAATTTAAAGAAATGTTTATACAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873
```

```
RESULT 27
US-10-847-983-13
; Sequence 13, Application US/10847983
```



```
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0.36
; US-10-847-983-13

Query Match 99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Qy 121 CTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATAC 180
Db 121 CTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATAC 180
Qy 181 ACTTTAAAAAATTTAAATTCCTGGAATTTGATGTAATCAATTTTATTTGAGCAATAC 240
Db 181 ACTTTAAAAAATTTAAATTCCTGGAATTTGATGTAATCAATTTTATTTGAGCAATAC 240
Qy 241 TACACCAAGCTCATCTAGAAATGCAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
Db 241 TACACCAAGCTCATCTAGAAATGCAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Db 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATCTTACAAAAATAAATCTATATGCTTATGCTCTTAATAGCAATTTT 660
Db 601 GAAAAAATCTTACAAAAATAAATCTATATGCTTATGCTCTTAATAGCAATTTT 660

; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0.36
; US-10-847-983-13

Query Match 99.3%; Score 869.8; DB 61; Length 873;
Best Local Similarity 99.8%; Pred. No. 9.3e-191;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Qy 121 CTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATAC 180
Db 121 CTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATAC 180
Qy 181 ACTTTAAAAAATTTAAATTCCTGGAATTTGATGTAATCAATTTTATTTGAGCAATAC 240
Db 181 ACTTTAAAAAATTTAAATTCCTGGAATTTGATGTAATCAATTTTATTTGAGCAATAC 240
Qy 241 TACACCAAGCTCATCTAGAAATGCAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
Db 241 TACACCAAGCTCATCTAGAAATGCAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Db 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATCTTACAAAAATAAATCTATATGCTTATGCTCTTAATAGCAATTTT 660
Db 601 GAAAAAATCTTACAAAAATAAATCTATATGCTTATGCTCTTAATAGCAATTTT 660
```

Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300  
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACCAACTTAAAGATTTAAAGCTTATTTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACCAACTTAAAGATTTAAAGCTTATTTTAAA 360  
Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420  
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Qy 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600  
Db 541 GATGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAAATTTATTTGCTTATGCTTAAATTTTATAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAAAATTTATTTGCTTATGCTTAAATTTTATAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATTAACACACT 720  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATTAACACACT 720  
Qy 721 AAGATATACCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780  
Db 721 AAGATATACCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780  
Qy 781 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAATTA 840  
Db 781 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAATTA 840

RESULT 29

US-10-850-807-13  
; Sequence 13, Application US/10850807  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,807  
; CURRENT FILING DATE: 2004-05-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-850-807-13

Query Match 99.3%; Score 869.8; DB 61; Length 873;  
Best Local Similarity 99.8%; Pred. No. 9.3e-191;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60  
Db 1 ATGAAAAAGTTATTTATTTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60  
Qy 61 CTACCAATGATTTTATTTGATGTTTATTTAGATGATTAATTTTATTTTGAAGATAAATCTAT 120  
Db 61 CTACCAATGATTTTATTTGATGTTTATTTAGATGATTAATTTTATTTTGAAGATAAATCTAT 120  
Qy 121 CTGTGTAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAAGCAATCTAC 180  
Db 121 CTGTGTAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAAGCAATCTAC 180  
Qy 181 ACTTTAAAAATTTTAAATCAAGAAATGAGACCGAACTAAATATATGTGTTCTAA 240  
Db 181 ACTTTAAAAATTTTAAATCAAGAAATGAGACCGAACTAAATATATGTGTTCTAA 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACCAATTTTAAAGAAATTTTAAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAATTTTAAAGAAATTTTAAATGCTTATTTTAAA 360  
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420  
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATA CAAGAAATTTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATA CAAGAAATTTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Qy 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600  
Db 541 GATGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTACGAAATTTT 660  
Db 601 GAAAAAATTTACAAAAATTTAAATTTTATGCTTATGCTTAAATAGTCTTTTACGAAATTTT 660  
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTTCAATTTTATCATACAGAAAAATTAACACACT 720  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTCAATTTTATCATACAGAAAAATTAACACACT 720  
Qy 721 AAGATATACCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780  
Db 721 AAGATATACCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780  
Qy 781 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAATTA 840  
Db 781 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAATTA 840  
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873

RESULT 30

US-10-961-882-13  
; Sequence 13, Application US/10961882  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada

;; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
;; FILE REFERENCE: Gangliosides and Ganglioside Mimics  
;; CURRENT APPLICATION NUMBER: US/10/961.882  
;; CURRENT FILING DATE: 2004-10-08  
;; PRIOR APPLICATION NUMBER: US/09/816,028  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/118,213  
;; PRIOR FILING DATE: 1999-02-01  
;; PRIOR APPLICATION NUMBER: US 09/495,406  
;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 13  
;; LENGTH: 873  
;; TYPE: DNA  
;; ORGANISM: Campylobacter jejuni  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(873)  
;; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
;; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-961-882-13

Query Match 99.3%; Score 869.8; DB 63; Length 873;  
Best Local Similarity 99.8%; Pred. No. 9.3e-191;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGCTTATTGCTGGAAATGGACCAAGTTTAAAGAAAAATGATTATTCAAGG 60  
DB 1 ATGAAAAAGCTTATTGCTGGAAATGGACCAAGTTTAAAGAAAAATGATTATTCAAGG 60

QY 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATCTAC 180  
DB 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATCTAC 180

QY 181 ACTTTAAAAATTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTTAAT 240  
DB 181 ACTTTAAAAATTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300

QY 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360  
DB 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360

QY 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTTGCGAGTAGCC 420  
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
DB 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAATCTTACAAAAATAAACTATATTGCTTTATGTCCTAATAGTCTTTTACGCAATTTT 660  
DB 601 GAAAAAATCTTACAAAAATAAACTATATTGCTTTATGTCCTAATAGTCTTTTACGCAATTTT 660

QY 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATATCATACAGAAAAATAACTACTAC 720

DB 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTAC 720

QY 721 AAAGATATCTATCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTAATTTT 780  
DB 721 AAAGATATCTATCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTAATTTT 780

QY 781 AAAAAATAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAAGATCTATTAAAGATTA 840  
DB 781 AAAAAATAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873  
DB 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 31  
US-10-962-235-13  
;; Sequence 13, Application US/10962235  
;; GENERAL INFORMATION:  
;; APPLICANT: Gilbert, Michel  
;; APPLICANT: Wakarchuk, Warren W.  
;; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
;; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
;; FILE REFERENCE: 019633-000111US  
;; CURRENT APPLICATION NUMBER: US/10/962,235  
;; CURRENT FILING DATE: 2004-10-08  
;; PRIOR APPLICATION NUMBER: US/09/816,028  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/118,213  
;; PRIOR FILING DATE: 1999-02-01  
;; PRIOR APPLICATION NUMBER: US 09/495,406  
;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 13  
;; LENGTH: 873  
;; TYPE: DNA  
;; ORGANISM: Campylobacter jejuni  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(873)  
;; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
;; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-962-235-13

Query Match 99.3%; Score 869.8; DB 63; Length 873;  
Best Local Similarity 99.8%; Pred. No. 9.3e-191;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGCTTATTGCTGGAAATGGACCAAGTTTAAAGAAAAATGATTATTCAAGG 60  
DB 1 ATGAAAAAGCTTATTGCTGGAAATGGACCAAGTTTAAAGAAAAATGATTATTCAAGG 60

QY 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATCTAC 180  
DB 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATCTAC 180

QY 181 ACTTTAAAAATTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTTAAT 240  
DB 181 ACTTTAAAAATTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCT 300

QY 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360

Db 301 GATGCTCAATGGGATAGATATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360  
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTGATACAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTGATACAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGACTATATCGGCATAGTAAATAACAGATATATAAGAGCTTTTAAAGATTTCTA 600  
Db 541 GATCGCTCGACTATATCGGCATAGTAAATAACAGATATATAAGAGCTTTTAAAGATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAACTAGTCTTTTAAAGAAATTTT 660  
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAACTAGTCTTTTAAAGAAATTTT 660  
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTTAACTATATGCTTATGCTTAACTAGTCTTTTAA 720  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTAACTATATGCTTATGCTTAACTAGTCTTTTAA 720  
Qy 721 AAAGATATCTATACCTCTAGTGAGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780  
Db 721 AAAGATATCTATACCTCTAGTGAGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780  
Qy 781 AAAAAATTTAAATTTAAAGAAATGTTTATTAAGATTTGATTAAGATTTTAAAGATTTA 840  
Db 781 AAAAAATTTAAATTTAAAGAAATGTTTATTAAGATTTGATTAAGATTTTAAAGATTTA 840  
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873

RESULT 32

US-10-962-334-13  
; Sequence 13, Application US/10962334  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,334  
; PRIOR FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (Cstii) from C. jejuni O:36  
US-10-962-334-13

Query Match 99.38; Score 869.8; DB 63; Length 873;  
Best Local Similarity 99.88; Pred. No. 9.3e-191;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTTATTTATTTCTGGAATGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTTATTTATTTCTGGAATGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60  
Qy 61 CTACAAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACAAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 121 CTTGTAAAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Db 121 CTTGTAAAAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTTAAT 240  
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTTAAT 240  
Qy 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTTATTTTCT 300  
Db 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTTATTTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Qy 361 TTTACGAAAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Db 361 TTTACGAAAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGACTATATCGGCATAGTAAATAACAGATATATAAGAGCTTTTAAAGATTTCTA 600  
Db 541 GATCGCTCGACTATATCGGCATAGTAAATAACAGATATATAAGAGCTTTTAAAGATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAACTAGTCTTTTAAAGAAATTTT 660  
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTAACTAGTCTTTTAAAGAAATTTT 660  
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTTAACTATATGCTTATGCTTAACTAGTCTTTTAA 720  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTTAACTATATGCTTATGCTTAACTAGTCTTTTAA 720  
Qy 721 AAAGATATCTATACCTCTAGTGAGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780  
Db 721 AAAGATATCTATACCTCTAGTGAGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780  
Qy 781 AAAAAATTTAAATTTAAAGAAATGTTTATTAAGATTTGATTAAGATTTTAAAGATTTA 840  
Db 781 AAAAAATTTAAATTTAAAGAAATGTTTATTAAGATTTGATTAAGATTTTAAAGATTTA 840  
Qy 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCATTTTCAAGGAAAA 873

RESULT 33

US-10-303-161-11  
; Sequence 11, Application US/10303161  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,161



Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTATGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTATGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAAATAAAACTATATGCTTATGCTTCTTATAGCAATTTT 660  
Db 601 GAAAAAATTTACAAAAATAAAACTATATGCTTATGCTTCTTATAGCAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAATTAACACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAATTAACACTACACT 720  
Qy 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTTAAATTTT 780  
Db 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTTAAATTTT 780  
Qy 781 AAAAAATAAAAATTAAGAAATTTTATTTATCAAGTTGATAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAAAATTAAGAAATTTTATTTATCAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

## RESULT 35

US-10-820-536-11  
; Sequence 11, Application US/10820536  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; PRIOR FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4

Query Match 98.9%; Score 866.6; DB 61; Length 873;  
Best Local Similarity 99.5%; Pred. No. 5.1e-190;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAAGTTATTTATTTGCTGGAAATGGCAAGTTTAAAAAGAAATTTGATTTCAAGG 60

Db 1 ATGAAAAAAGTTATTTATTTGCTGGAAATGGCAAGTTTAAAAAGAAATTTGATTTCAAGG 60  
Qy 61 CTACCAATGATTTTGTATTTTATAGATGTAATCAATTTTATTTTGAAGATATAACTAT 120  
Db 61 CTACCAATGATTTTGTATTTTATAGATGTAATCAATTTTATTTTGAAGATATAACTAT 120  
Qy 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTCTTTGAGCAATACTAC 180  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACCTAATTTATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACCTAATTTATGTTCTTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAATCGAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAATCGAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA 360  
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420  
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAATTAACACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAATTAACACTACACT 720  
Qy 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTTAAATTTT 780  
Db 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTTAAATTTT 780  
Qy 781 AAAAAATAAAAATTAAGAAATTTTATTTATCAAGTTGATAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAAAATTAAGAAATTTTATTTATCAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

## RESULT 36

US-10-821-573-11  
; Sequence 11, Application US/10821573  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,573  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128

```

; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-821-573-11

Query Match      98.9%; Score 866.6; DB 61; Length 873;
Best Local Similarity 99.5%; Pred. No. 5.1e-190;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
DB 1 ATGAAAAAGTTATTATGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

QY 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAATACTAT 120
DB 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAATACTAT 120

QY 121 CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
DB 121 CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTGGTTCCTTCTTGAGCAATACTAC 180

QY 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGCCGAACCTAATTTGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGCCGAACCTAATTTGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTTTTCCT 300
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTGTAAAAACCTTTTAAAGAAATTTAATGCTTTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTGTAAAAACCTTTTAAAGAAATTTAATGCTTTTAAA 360

QY 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
DB 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAATCTTACAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACGAAATTTT 660
DB 601 GAAAAAATCTTACAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACGAAATTTT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATCAAGAAAAATAAATCTACACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATCAAGAAAAATAAATCTACACT 720

QY 721 AAAGATATACTCATCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
DB 721 AAAGATATACTCATCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780

```

361	Qy	TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAGTAGCC	420
361	Db	TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAGTAGCC	420
421	Qy	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
421	Db	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
481	Qy	TCATTATGCTTTTGATATACAAACAAAGAAATCTCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
481	Db	TCATTATGCTTTTGATATACAAACAAAGAAATCTCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
541	Qy	GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
541	Db	GATCGCTCACATATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
601	Qy	GAAAAAATTACAAAAATAAACTATAATGCTTTATGTCCTATAAGTCTTTTAGCAAAATTTT	660
601	Db	GAAAAAATTACAAAAATAAACTATAATGCTTTATGTCCTATAAGTCTTTTAGCAAAATTTT	660
661	Qy	ATAGAACTAGGCCAAATTTTAAATTCAAATTTTATCACAAGAAAAATAAATCACTACT	720
661	Db	ATAGAACTAGGCCAAATTTTAAATTCAAATTTTATCACAAGAAAAATAAATCACTACT	720
721	Qy	AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATAATTAATTTT	780
721	Db	AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATAATTAATTTT	780
781	Qy	AAAAAATAAAAAATTTAAAGAAAAATGTTTATTTACAAGTTTGATAAAAAAGATCTATTAAGATTA	840
781	Db	AAAAAATAAAAAATTTAAAGAAAAATGTTTATTTACAAGTTTGATAAAAAAGATCTATTAAGATTA	840
841	Qy	CTTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
841	Db	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873

```

RESULT 38
US-10-830-825-11
; Sequence 11, Application US/10830825
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011US
; CURRENT APPLICATION NUMBER: US/10/830,825
; CURRENT FILING DATE: 2004-04-24
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-830-825-11

```

RESULT 39  
US-10-830-997-11  
; Sequence 11, Application US/10830997  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-00011US  
; CURRENT APPLICATION NUMBER: US/10/830,997  
; CURRENT FILING DATE: 2004-04-22  
; PRIOR APPLICATION NUMBER: US/10/303,128

TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 FILE REFERENCE: 019633-000111US  
 CURRENT APPLICATION NUMBER: US/10/830,997  
 CURRENT FILING DATE: 2004-04-22  
 PRIOR APPLICATION NUMBER: US/10/303,128



; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-830-997-11

Query Match 98.9%; Score 866.6; DB 61; Length 873;
Best Local Similarity 99.5%; Pred. No. 5.1e-190;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTTGACATATAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGTTCTTTTGACATATAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATTTATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTAAA 360

Qy 361 TTTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGACATATACGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACATATACGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 601 GAAAAAATTTACAAAAATAAAAATCTATTTGCTTATGCTCCTAATAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAAAATCTATTTGCTTATGCTCCTAATAGTCTTTTAGCAAAATTT 660

Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATATACAGAAAAAATAAATACTAC 720
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATATACAGAAAAAATAAATACTAC 720

Qy 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780

Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780
Qy 781 AAAAAAATAAAAAATAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAAGACATTTATTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAAGACATTTATTCAAAGGAAAA 873

RESULT 40
US-10-845-408-11
; Sequence 11. Application US/10845408
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: Gangliosides and Ganglioside Mimics
; CURRENT APPLICATION NUMBER: US/10/845,408
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-845-408-11

Query Match 98.9%; Score 866.6; DB 61; Length 873;
Best Local Similarity 99.5%; Pred. No. 5.1e-190;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTTGACATATAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGTTCTTTTGACATATAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAAATTTATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTAATTTAAA 360

Qy 361 TTTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420

Db 361 TTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
Qy 541 GATCCTCGACTATATCGGCATATAGTAAATATACAGATATAAAGCTTTAGAAATCTA 600  
Db 541 GATCCTCGACTATATCGGCATATAGTAAATATACAGATATAAAGCTTTAGAAATCTA 600  
Qy 601 GAAAAAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 660  
Db 601 GAAAAAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATTAACACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATTAACACACT 720  
Qy 721 AAAGATATACATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAATTAATTAATTT 780  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAATTAATTAATTT 780  
Qy 781 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840  
Db 781 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

RESULT 41  
US-10-845-412-11  
; Sequence 11, Application US/10845412  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/845,412  
; PRIORITY FILING DATE: 2004-05-12  
; PRIORITY APPLICATION NUMBER: US/10/303,128  
; PRIORITY FILING DATE: 2002-11-21  
; PRIORITY FILING DATE: 2001-03-21  
; PRIORITY FILING DATE: 1999-02-01  
; PRIORITY FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-845-412-11  
Query Match 98.9%; Score 866.6; DB 61; Length 873;  
Best Local Similarity 99.5%; Pred. No. 5,1e-190;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTCAAGG 60

Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTCAAGG 60  
Qy 61 CTACCAAAATGATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTACCAAAATGATTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Qy 121 CTGTTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTTCTTTGAGCAATACATAC 180  
Db 121 CTGTTAAAAAATGCAAGCAGTGTGTTTACACCCCTGGTTTCTTTCTTTGAGCAATACATAC 180  
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATATGTTGTTCTTAAT 240  
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATATATGTTGTTCTTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTTGCGAGTAGCC 420  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAAGAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAAGAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAAT 540  
Qy 541 GATCGCTCGCACTATATCGGCATATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGCATATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 660  
Db 601 GAAAAAATTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATTAACACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATTAACACACT 720  
Qy 721 AAAGATATATCTATACCTTCTAGTGAGGCTTTATGAAAAATTTTCAAAATTAATTAATTT 780  
Db 721 AAAGATATATCTATACCTTCTAGTGAGGCTTTATGAAAAATTTTCAAAATTAATTAATTT 780  
Qy 781 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840  
Db 781 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

RESULT 42  
US-10-846-219-11  
; Sequence 11, Application US/10846219  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/846,219  
; PRIORITY FILING DATE: 2004-05-14  
; PRIORITY APPLICATION NUMBER: US/09/816,028

```

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-846-219-11

Query Match      98.9%; Score 866.6; DB 61; Length 873;
Best Local Similarity 99.5%; Pred. No. 5.1e-190;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60

Qy 61 CTACCAATGATTTTGATGTTTATGATGTTATCAATTTTATTTGAAGATAATACTAT 120
Db 61 CTACCAATGATTTTGATGTTTATGATGTTATCAATTTTATTTGAAGATAATACTAT 120

Qy 121 CTTCGTTAAAAATGCAAGCAGTGTTCCTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTCGTTAAAAATGCAAGCAGTGTTCCTACACCCCTGTTTCTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy 241 TACAAACAGCTCATCTAGAAAAATGAAAAATTTGTAAAAACCTTTTACGATATTATTCCT 300
Db 241 TACAAACAGCTCATCTAGAAAAATGAAAAATTTGTAAAAACCTTTTACGATATTATTCCT 300

Qy 301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAT 360
Db 301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAT 360

Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600

Qy 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGTCCTAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGTCCTAATAGTCTTTTACGAAATTTT 660

Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATTAACCTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATTAACCTACACT 720

Qy 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780

```

```

781 AAAAAAATAAAAAATTAAGAAAAATCTTTATTACAAGTTGATAAAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTAAGAAAAATCTTTATTACAAGTTGATAAAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 43
US-10-847-983-11
; Sequence 11, Application US/10847983
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-847-983-11

Query Match      98.9%; Score 866.6; DB 61; Length 873;
Best Local Similarity 99.5%; Pred. No. 5.1e-190;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60

Qy 61 CTACCAATGATTTTGATGTTTATGATGTTATCAATTTTATTTGAAGATAATACTAT 120
Db 61 CTACCAATGATTTTGATGTTTATGATGTTATCAATTTTATTTGAAGATAATACTAT 120

Qy 121 CTTCGTTAAAAATGCAAGCAGTGTTCCTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTCGTTAAAAATGCAAGCAGTGTTCCTACACCCCTGTTTCTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy 241 TACAAACAGCTCATCTAGAAAAATGAAAAATTTGTAAAAACCTTTTACGATATTATTCCT 300
Db 241 TACAAACAGCTCATCTAGAAAAATGAAAAATTTGTAAAAACCTTTTACGATATTATTCCT 300

Qy 301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAT 360
Db 301 GATCCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAT 360

Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600

Qy 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGTCCTAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGTCCTAATAGTCTTTTACGAAATTTT 660

Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATTAACCTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATTAACCTACACT 720

Qy 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATCTCATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780

```

Qy	421	ATAGCCCTAGGATCAAAAGAAATTTATCTTTCGGGAATTTGATTTTTATCAAAATCGGTCA	480
Db	421	ATAGCCCTAGGATCAAAAGAAATTTATCTTTCGGGAATTTGATTTTTATCAAAATCGGTCA	480
Qy	481	TCATTATGCTTTTGATACCAACAAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Db	481	TCATTATGCTTTTGATACCAACAAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Qy	541	GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
Db	541	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
Qy	601	GAAAAAACTTCAAAAAATAAAACTATATTCCTATGTCCTAACTAGTCTTTTACGAATTTT	660
Db	601	GAAAAAACTTCAAAAAATAAAACTATATTCCTATGTCCTAACTAGTCTTTTACGAATTTT	660
Qy	661	ATGAACTAGCGCCAAATTTTAAATTTCAAATTTTATCATACAAGAAAAAAATAACTACACT	720
Db	661	ATGAACTAGCGCCAAATTTTAAATTTCAAATTTTATCATACAAGAAAAAAATAACTACACT	720
Qy	721	AAAGATATCTCATACCTTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT	780
Db	721	AAAGATATCTCATACCTTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT	780
Qy	781	AAAAAATAAAAATTTAAAGAAAAATGTTTATTACAAGTTGATAAAAAGATCTATTAAGATTA	840
Db	781	AAAAAATAAAAATTTAAAGAAAAATGTTTATTACAAGTTGATAAAAAGATCTATTAAGATTA	840
Qy	841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
Db	841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873

```

RESULT 44
US-10-850-125-11
; Sequence 11, Application US/10850125
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-850-125-11

```

RESULT 45

```

US-10-850-807-11
; Sequence 11, Application US/10850807
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406

```

;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 11  
;; LENGTH: 873  
;; TYPE: DNA  
;; ORGANISM: Campylobacter jejuni  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(873)  
;; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
;; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-850-807-11

Query Match 98.9%; Score 866.6; DB 61; Length 873;  
Best Local Similarity 99.5%; Pred. No. 5.1e-190;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGG 60

Qy 61 CTACCAATGATTTTGATGTTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGATGTTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAATGCAAGAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATGCAAGAGTGTTTTACACCCCTGTTTCTTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTGTTCTAAT 240  
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300

Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360

Qy 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420  
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTTAGAAATTTCTA 600

Qy 601 GAAAAAATCTTACAAAATAAAAACTATTTGCTTATGTCTCTAATAGTCTTTTACGAAATTTT 660  
Db 601 GAAAAAATCTTACAAAATAAAAACTATTTGCTTATGTCTCTAATAGTCTTTTACGAAATTTT 660

Qy 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATAAATACTACT 720  
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATAAATACTACT 720

Qy 721 AAAGATATATCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATATCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780

Qy 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTTACAGTTGATAAAGATCTATTAGATT 840  
Db 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTTACAGTTGATAAAGATCTATTAGATT 840

Qy 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 46  
US-10-961-882-11  
; Sequence 11, Application US/10961882  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/961,882  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-961-882-11

Query Match 98.9%; Score 866.6; DB 63; Length 873;  
Best Local Similarity 99.5%; Pred. No. 5.1e-190;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGG 60

Qy 61 CTACCAATGATTTTGATGTTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGATGTTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAATGCAAGAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATGCAAGAGTGTTTTACACCCCTGTTTCTTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTGTTCTAAT 240  
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300

Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360

Qy 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420  
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480

Qy	481	TC	TATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Db	481	TC	TATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Qy	541	GATCGCTCGCACTATATCGGACATAGTAAAAAATACAGATATAAAAGCTTTTAGAATTTCTTA	600	
Db	541	GATCGCTCACACTATATCGGACATAGTAAAAAATACAGATATAAAAGCTTTTAGAATTTCTTA	600	
Qy	601	GA	AAAAAATTTACAAAATAAAAATATATTTGCTTATGTCCTTAATAGTCTTTTTCAGCAAAATTTT	660
Db	601	GA	AAAAAATTTACAAAATAAAAATATATTTGCTTATGTCCTTAACAGTCTTTTTCAGCAAAATTTT	660
Qy	661	ATGA	AACTAGGCCCAAAATTTTAAATTCAAATTTTATCATACAAGAAAAAAAATAACTACACT	720
Db	661	ATGA	AACTAGGCCCAAAATTTTAAATTCAAATTTTATCATACAAGAAAAAAAATAACTACACT	720
Qy	721	AA	GATATACCTCATACCTCTTCAGTGAGGCTTATGGAAAAATTTTCAAAAAAATATTAATTTT	780
Db	721	AA	GATATACCTCATACCTCTTCAGTGAGGCTTATGGAAAAATTTTCAAAAAAATATTAATTTT	780
Qy	781	AAAA	AAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTTGATAAAAAGACTTATTAAGATTTA	840
Db	781	AAAA	AAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTTGATAAAAAGACTTATTAAGATTTA	840
Qy	841	CT	TAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
Db	841	CT	TAGTGATATAAAGCATTTATTTCAAAGGAAAA	873

```

RESULT 47
US-10-962-235-11
; Sequence 11, Application US/10962235
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011US
; CURRENT APPLICATION NUMBER: US/10/962,235
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816, 028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118, 213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495, 406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-962-235-11

```

[illegible]

```

RESULT 48
US-10-962-334-11
; Sequence 11, Application US/109622334
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/962,334
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11

```

```
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-962-334-11

Query Match      98.9%; Score 866.6; DB 63; Length 873;
Best Local Similarity 99.5%; Pred. No. 5.1e-130;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60

Qy 61 CTACCAATGATTTTGATGTAATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGTAATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGACCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTGGTTCTTCTTTGACCAATACTAC 180

Qy 181 ACTTTAAACAATTTAAATCCAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACAATTTAAATCCAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTGTAACCTTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTGTAACCTTTTACGATTTATTTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360

Qy 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTCTGTAACCTTTTACGATTTATTTTCCCT 420
Db 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTCTGTAACCTTTTACGATTTATTTTCCCT 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600

Qy 601 GAAAAAATTTACAAAATAAACCTATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAACCTATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660

Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATATCAAGAAAAAATAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATATCAAGAAAAAATAAATACTACACT 720

Qy 721 AAAGATATCTCATACCTTCTAGTGAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTCATACCTTCTAGTGAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTTT 780

Qy 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTATTAAGCAATTTTCAAGGAAA 873
Db 841 CCTAGTATTAAGCAATTTTCAAGGAAA 873
```

## RESULT 49

```
US-10-303-161-4
; Sequence 4, Application US/10303161
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-161-4
```

```
Query Match      96.3%; Score 844; DB 47; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60

Qy 61 CTACCAATGATTTTGATGTAATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGTAATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGACCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCTCTGGTCTTTTGTGAACAATACTAC 180

Qy 181 ACTTTAAACAATTTAAATCCAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACAATTTAAATCCAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAATGCTTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAATGCTTTATTTTAAA 360

Qy 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTCAATCAAGAAATTTTATATGTGTCAGTAGCC 420
Db 361 TTTTACCAAGCTCATCTAGAAAATGAAAATTTTCAATCAAGAAATTTTATATGTGTCAGTAGCT 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAAT 540
```

```
Db 481 TCTTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAACTGGCTCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGCATAGTAAATAATACAGATATATAAAGCTTTTAGAAATTCTTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAATAATACAGATATATAAAGCTTTTAGAAATTCTTA 600
Qy 601 GAAAAAATCTTACAAATAAATATATGCTTTATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATCTTACAAATAAATATATGCTTTATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAAATAACTACACT 720
Qy 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAAAATATTAAATTTT 780
Db 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAAAATATTAAATTTT 780
Qy 781 AAAAAATATAAATAAATAAAGAAAAATCTTTATTACAAGTTGATAAAGATCTATTAAAGATT 840
Db 781 AAAAAATATAAATAAATAAAGAAAAATCTTTATTACAAGTTGATAAAGATCTATTAAAGATT 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

RESULT 50
US-10-734-719-4
; Sequence 4, Application US/10734719
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/734,719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-734-719-4

Query Match 96.3%; Score 844; DB 60; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAAAGAAATGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAAAGAAATGATTTTCAAGG 60

Qy 61 CTACCAATGATTTTATGATGATTTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTATGATGATTTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTGTGTAATAAATGCAAGACAGTGTTTACACCCCTAAATTTCTTCTTGAGCAACTACTAC 180
```

```
Db 121 CTGTGTAATAAATGCAAGACAGTGTTTACCAATCTCTGGTCTTTTGTGAACAATACTACTAC 180
Qy 181 ACTTTTAAACAATTTAATCCAAATCAAGAAATAGAGCCGAACCTAATATATGTTCTTAAT 240
Db 181 ACTTTTAAACAATTTAATCCAAATCAAGAAATAGAGCCGAACCTAATATATGTTCTTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATTTATTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATTTATTTCCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 361 TTTTCAGGAATTTTCAATCAAGAAATTTACCTCAGGGCTCTATATGCTGAGTAGGCC 420
Db 361 TTTTCAGGAATTTTCAATCAAGAAATTTACCTCAGGGCTCTATATGCTGAGTAGGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGCATAGTAAATAATACAGATATATAAAGCTTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATAATACAGATATATAAAGCTTTTAGAAATTTCTA 600
Qy 601 GAAAAAATCTTACAAAAATAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATCTTACAAAAATAAATCTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAAATAACTACACT 720
Qy 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780
Db 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780
Qy 781 AAAAAATATAAATAAATAAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATT 840
Db 781 AAAAAATATAAATAAATAAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATT 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

RESULT 51
US-10-820-536-4
; Sequence 4, Application US/10820536
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
```





Qy 421 ATAGCCCTGATGATACAAAGAAATTAATCTTTGGGAATGATTTTATCAAAAATGGGTCA 480  
Db |||||  
Qy 421 ATAGCCCTGATGATACAAAGAAATTAATCTTTGGGAATGATTTTATCAAAAATGGGTCA 480  
Db |||||  
Qy 481 TCTTATGCTTTGATACCAAAACAAAGAAATCTTTTAAAACTGAGCCCTGATTTTAAAAAT 540  
Db |||||  
Qy 541 GATCGCTGCTATATCGGCATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600  
Db |||||  
Qy 601 GAAAAAATCAAAAATAAAATATATGCTTATGCTCTTAATAGTCTTTTACGAAATTTT 660  
Db |||||  
Qy 601 GAAAAAATCAAAAATAAAATATATGCTTATGCTCTTAATAGTCTTTTACGAAATTTT 660  
Db |||||  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATPACACT 720  
Db |||||  
Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATAATTTT 780  
Db |||||  
Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATAATTTT 780  
Db |||||  
Qy 781 AAAAAAATAAAATTAAGAAATGCTTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Db |||||  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
Db |||||  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
Db |||||

RESULT 53

US-10-821-604-4  
; Sequence 4, Application US/10821604  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
; OTHER INFORMATION: biosynthesis locus)  
US-10-821-604-4  
Query Match 96.3%; Score 844; DB 61; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8,9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGCTTTTAAAGAAATGATTTATTTCAAGG 60

Db |||||  
Qy 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGCTTTTAAAGAAATGATTTATTTCAAGG 60  
Db |||||  
Qy 61 CTACCAAAATGATTTTATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db |||||  
Qy 61 CTACCAAAATGATTTTATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db |||||  
Qy 121 CTTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180  
Db |||||  
Qy 121 CTTGTGTAATAAATGCAAAAGCAGTGTTTTACCAATCTCTGGCTTTTGTGAACAATACTAC 180  
Db |||||  
Qy 181 ACTTTAAAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATATGTTCTTAAT 240  
Db |||||  
Qy 181 ACTTTAAAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATATGTTCTTAAT 240  
Db |||||  
Qy 241 TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAAACTTTTACGATTAATTTTCCCT 300  
Db |||||  
Qy 241 TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAAACTTTTACGATTAATTTTCCCT 300  
Db |||||  
Qy 301 GATGCTCATTTCTGGGATATGATTTTAAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db |||||  
Qy 301 GATGCTCATTTCTGGGATATGATTTTAAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360  
Db |||||  
Qy 361 TTTTCAGGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Db |||||  
Qy 361 TTTTCAGGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCT 420  
Db |||||  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGGAAATGATTTTATCAAAAATGGGTCA 480  
Db |||||  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGGAAATGATTTTATCAAAAATGGGTCA 480  
Db |||||  
Qy 481 TCTTATGCTTTTATGATACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db |||||  
Qy 481 TCTTATGCTTTTATGATACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db |||||  
Qy 541 GATCGCTCGCATCTATATCGGCATAGTAAAAATACAGATATAAAAAGCTTTTAGAAATTTCTA 600  
Db |||||  
Qy 541 GATCGCTCGCATCTATATCGGCATAGTAAAAATACAGATATAAAAAGCTTTTAGAAATTTCTA 600  
Db |||||  
Qy 601 GAAAAAATCAAAAATAAAATCTATATGCTTATGCTTAAATAGTCTTTTATAGCAAAATTTT 660  
Db |||||  
Qy 601 GAAAAAATCAAAAATAAAATCTATATGCTTATGCTTAAATAGTCTTTTATAGCAAAATTTT 660  
Db |||||  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATPACACT 720  
Db |||||  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATPACACT 720  
Db |||||  
Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATAATTTT 780  
Db |||||  
Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATAATTTT 780  
Db |||||  
Qy 781 AAAAAAATAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db |||||  
Qy 781 AAAAAAATAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db |||||  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
Db |||||  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
Db |||||

RESULT 54

US-10-830-825-4  
; Sequence 4, Application US/10830825  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/830,825  
; CURRENT FILING DATE: 2004-04-24  
; PRIOR APPLICATION NUMBER: US/09/816,028

```
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-830-825-4

Query Match          96.3%; Score 844; DB 61; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTCTTTGGAGCAATAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACAACTCTGCTCTTTTGTGAACAATACTAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAAATATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAAATATGTTCTTAAT 240

Qy 241 TACACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTTPAAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTPAAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360

Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCT 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTCGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTCGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAAAATCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAAAATCTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600

Qy 601 GAAAAAATCTTACAAAATAAACATATTTGCTTATGCTTCTTAATAGCTTTTGAAGAAATTTT 660
Db 601 GAAAAAATCTTACAAAATAAACATATTTGCTTATGCTTCTTAATAGCTTTTGAAGAAATTTT 660

Qy 661 ATAGAACTAGCGGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACACT 720
Db 661 ATAGAACTAGCGGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACACT 720

Qy 721 AAAGATATATCTATCTTCTAGTGGGCTTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
```

```
Db 721 AAAGATATATCTATCTTCTAGTGGGCTTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAGATTAA 840
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAGATTAA 840
Qy 841 CCTAGTGATATAAAGCATTATTTCAAAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAAGCATTATTTCAAAAGGAAAAATAA 876

RESULT 55
US-10-830-997-4
; Sequence 4, Application US/10830997
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/830,997
; CURRENT FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-830-997-4

Query Match          96.3%; Score 844; DB 61; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGCTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy 61 CTACCAAAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTCTTTGGAGCAATAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACAACTCTGCTCTTTTGTGAACAATACTAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAAATATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAAATATGTTCTTAAT 240

Qy 241 TACACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTTCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTTPAAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTPAAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360

Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCT 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTCGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTCGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAAAATCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAAAATCTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600

Qy 601 GAAAAAATCTTACAAAATAAACATATTTGCTTATGCTTCTTAATAGCTTTTGAAGAAATTTT 660
Db 601 GAAAAAATCTTACAAAATAAACATATTTGCTTATGCTTCTTAATAGCTTTTGAAGAAATTTT 660

Qy 661 ATAGAACTAGCGGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACACT 720
Db 661 ATAGAACTAGCGGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACACT 720

Qy 721 AAAGATATATCTATCTTCTAGTGGGCTTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
```

Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360  
QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCT 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTTAAACACTAGCCCTGATTTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTTAAACACTGGCTCCTGATTTTTAAAAAT 540  
QY 541 GATCCTCGCACTATATCGGACATAGTAAAGATACAGATATAAGAGCTTTAGAAATTTCTA 600  
Db 541 GATCCTCGCACTATATCGGACATAGTAAAGATACAGATATAAGAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAATGCTTTTACCAAAATTTT 660  
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAATGCTTTTACCAAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
QY 721 AAAGATATACATACACTTTAGTGAGGCTTTAGGAAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATACATACACTTTAGTGAGGCTTTAGGAAATTTTCAAAAAATTTAATTTT 780  
QY 781 AAAAAATTAATAATTAAGAAATCTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATTAATAATTAAGAAATCTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876

RESULT 56  
US-10-845-408-4  
; Sequence 4, Application US/10845408  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/845,408  
; CURRENT FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
; OTHER INFORMATION: biosynthesis locus  
US-10-845-408-4

Query Match 96.3%; Score 844; DB 61; Length 876;

Best Local Similarity 97.7%; Pred. No. 8,9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATCAAAAAAGTTATTTATTTCTGGAATGGACCAAGTTTAAAGAAAAATTTGATTTTCAAGG 60  
Db 1 ATCAAAAAAGTTATTTATTTCTGGAATGGACCAAGTTTAAAGAAAAATTTGATTTTCAAGG 60  
QY 61 CTACCAAAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAAATCTAT 120  
Db 61 CTACCAAAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAAATCTAT 120  
QY 121 CTTGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATATCTAC 180  
Db 121 CTTGTAAAAATGCAAAAGCAGTGTTTTACACCCCTCTTTTAAATCTCTTTTGTAGCAATATCTAC 180  
QY 181 ACTTTAAAAATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTGTCTTAAT 240  
Db 181 ACTTTAAAAATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTGTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360  
QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTTCAGAAATTTATCTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCT 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACACTAGCCCTGATTTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACACTGGCTCCTGATTTTTAAAAAT 540  
QY 541 GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGACATATATCGGACATAGTAAAAATACAGATATAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAACAGCTCTTTTAGCAAAATTT 660  
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAACAGCTCTTTTAGCAAAATTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
QY 721 AAAGATATACATACACTTTCTAGTGAGGCTTTAGGAAAAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATACATACACTTTCTAGTGAGGCTTTAGGAAAAATTTTCAAAAAATTTAATTTT 780  
QY 781 AAAAAATTAATAATTAAGAAATCTTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATTAATAATTAAGAAATCTTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876

RESULT 57  
US-10-845-412-4  
; Sequence 4, Application US/10845412  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics

FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/845,412  
CURRENT FILING DATE: 2004-05-12  
PRIOR APPLICATION NUMBER: US/10/303,128  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (CstII) from C. jejuni  
OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
OTHER INFORMATION: biosynthesis locus)  
US-10-845-412-4

Query Match 96.3%; Score 844; DB 61; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTGCTGGAATCGCAAGTTTAAAAAGAAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTGCTGGAATCGCAAGTTTAAAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAACTAT 120  
QY 121 CTGTTAAAAATGCAAGAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
DB 121 CTGTTAAAAATGCAAGAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATGAGCCGAACCTAATTTATGTTCTAAT 240  
DB 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATGAGCCGAACCTAATTTATGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
QY 361 TTTTACCAAGATTTTCAATCAAGAAATTAATGCTGAGGCTATATGTCGAGTAGCT 420  
DB 361 TTTTACCAAGATTTTCAATCAAGAAATTAATGCTGAGGCTATATGTCGAGTAGCT 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAATTTGATTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACAAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACAAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGACTATATCGGCATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGACTATATCGGCATAGTAAATATACAGATATAAAGCTTTTAGAAATTTCTA 600  
QY 601 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
DB 601 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660

QY 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACACT 720  
DB 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTACACT 720  
QY 721 AAAGATATACATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
DB 721 AAAGATATACATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAATAAAAATTAAGAAATGTTTATTTACAGTTGATTAAGAGATCTATTAAAGATTA 840  
DB 781 AAAAAATAAAAATTAAGAAATGTTTATTTACAGTTGATTAAGAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTTATTTCAGAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAGCATTTATTTCAGAGGAAAAATAA 876

## RESULT 58

US-10-846-219-4  
Sequence 4, Application US/10846219  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
TITLE OF INVENTION: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/846,219  
PRIOR FILING DATE: 2004-05-14  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (CstII) from C. jejuni  
OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
OTHER INFORMATION: biosynthesis locus)  
US-10-846-219-4

Query Match 96.3%; Score 844; DB 61; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTGCTGGAATCGCAAGTTTAAAAAGAAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTGCTGGAATCGCAAGTTTAAAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGTGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGTGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTTAAAAATGCAAGAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
DB 121 CTGTTAAAAATGCAAGAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATGAGCCGAACCTAATTTATGTTCTAAT 240  
DB 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATGAGCCGAACCTAATTTATGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300

Db 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAGATTTAAATGCTTATTTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTAA 360  
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTGTCAGTAGCT 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCCTCGCAGTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTACAAATTTCTA 600  
Db 541 GATCCTCGCAGTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTACAAATTTCTA 600  
Qy 601 GAAAAAATTTTACAAAAATATAATATGCTTATGTCCTTAATAGTCTTTTACCAATTTT 660  
Db 601 GAAAAAATTTTACAAAAATATAATATGCTTATGTCCTTAATAGTCTTTTACCAATTTT 660  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTAAATTTTATCATACAAAGAAAAATTAACACT 720  
Db 661 ATAGAACTAGCCCAAAATTTAAATTTAAATTTTATCATACAAAGAAAAATTAACACT 720  
Qy 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAAATTAATTTT 780  
Db 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAAATTAATTTT 780  
Qy 781 AAAAAATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

RESULT 59  
US-10-847-983-4  
; Sequence 4, Application US/10847983  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/847,983  
; PRIORITY FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; NAME/KEY: CDS  
; LOCATION: (1)...(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))

; OTHER INFORMATION: biosynthesis locus)  
US-10-847-983-4  
Query Match 96.3%; Score 844; DB 61; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAAAATGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAAAATGATTATTTCAAGG 60  
Qy 61 CTACCAAAATGATTTTGATGTAATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGATGTAATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120  
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATTTCAAGCAGTATTTTACAATCTCTGGTCTTTTGTGAACTACTAC 180  
Qy 181 ACTTTAAAAATTTAATCCAAAATCAAGATATGAGACCGAACTAAATTTATGTGTCTTAAT 240  
Db 181 ACTTTAAAAATTTAATCCAAAATCAAGATATGAGACCGAACTAAATTTATGTGTCTTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATTTTACGATTTATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATTTTACGATTTATTTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCC 420  
Db 361 TTTTACGAAATTTATCTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCT 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCAGTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCGCAGTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
Qy 601 GAAAAAATTTTACAAAAATAAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTT 660  
Db 601 GAAAAAATTTTACAAAAATAAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTT 660  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTTAAATTTTATCATACAAAGAAAAATTAACACT 720  
Db 661 ATAGAACTAGCCCAAAATTTAAATTTTAAATTTTATCATACAAAGAAAAATTAACACT 720  
Qy 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAAATTAATTTT 780  
Db 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAAATTAATTTT 780  
Qy 781 AAAAAATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876

RESULT 60  
US-10-850-125-4  
; Sequence 4, Application US/10850125  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel

```
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850.125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-850-125-4

Query Match          96.3%; Score 844; DB 61; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTATTATTGCTGGAAATGAGCAAGTTTAAAGAAATGATTATTCAGG 60
DB 1 ATGAAAAAGTATTATTGCTGGAAATGAGCAAGTTTAAAGAAATGATTATTCAGG 60
QY 61 CTACCAATGATTTTGATGTTATGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
DB 61 CTACCAATGATTTTGATGTTATGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
QY 121 CTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180
DB 121 CTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAACTAAATATGTTCTTAAT 240
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAACTAAATATGTTCTTAAT 240
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTTCT 300
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
QY 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCAAGCC 420
DB 361 TTTCAAGAAATTTATCTCAATCAAGAAATTTACCTCAGGAGTCTATATGTTGCAAGCT 420
QY 421 ATAGCCCTAGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGTCA 480
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAATGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTGCATATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAAGCTTACAAAATATAATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
```

```
DB 601 GAAAAAAGCTTACAAAATATAATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACT 720
DB 661 ATAGAACTAGCCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
DB 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAAATAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTTATAAGATTA 840
DB 781 AAAAAAATAAAAATTAAGAAAAATATTTTATTAACAAGTTGATAAAGATCTTATAAGATTA 840
QY 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876
DB 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876
```

## RESULT 61

```
US-10-850-807-4
; Sequence 4, Application US/10850807
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-850-807-4
```

```
Query Match          96.3%; Score 844; DB 61; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAGTATTATTGCTGGAAATGAGCAAGTTTAAAGAAATGATTATTCAGG 60
DB 1 ATGAAAAAGTATTATTGCTGGAAATGAGCAAGTTTAAAGAAATGATTATTCAGG 60
QY 61 CTACCAATGATTTTGATGTTATGATGTAATCAATTTTATTTGAGCAATACTAT 120
DB 61 CTACCAATGATTTTGATGTTATGATGTAATCAATTTTATTTGAGCAATACTAT 120
QY 121 CTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
DB 121 CTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAACTAAATATGTTCTTAAT 240
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAACTAAATATGTTCTTAAT 240
```



```
QY 241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAACACTTTTACGATTTATTTCT 300
DB |||||||
QY 241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAACACTTTTACGATTTATTTCT 300
DB |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
DB |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
DB |||||||
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTATATGTCAGTAGCT 420
DB |||||||
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTATATGTCAGTAGCT 420
DB |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480
DB |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480
DB |||||||
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
DB |||||||
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
DB |||||||
QY 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
DB |||||||
QY 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
DB |||||||
QY 601 GAAAAAATTTTAAAGAAATTTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 660
DB |||||||
QY 601 GAAAAAATTTTAAAGAAATTTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 660
DB |||||||
QY 661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 720
DB |||||||
QY 661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 720
DB |||||||
QY 721 AAAGATATCTATACCTCTTAGTGGGCTTATGGAATTTTCAAAATTTTAAATTTT 780
DB |||||||
QY 721 AAAGATATCTATACCTCTTAGTGGGCTTATGGAATTTTCAAAATTTTAAATTTT 780
DB |||||||
QY 781 AAAAAATTTAAAGAAATTTTAAAGAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTT 840
DB |||||||
QY 781 AAAAAATTTAAAGAAATTTTAAAGAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTT 840
DB |||||||
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAATATA 876
DB |||||||
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAATATA 876
DB |||||||
```

## RESULT 62

US-10-961-882-4

; Sequence 4, Application US/10961882

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-0001105

; CURRENT APPLICATION NUMBER: US/10/961,882

; PRIORITY FILING DATE: 2004-10-08

; PRIORITY FILING DATE: US/09/816,028

; PRIORITY FILING DATE: 2001-03-21

; PRIORITY FILING DATE: US 60/118,213

; PRIORITY FILING DATE: 1999-02-01

; PRIORITY FILING DATE: US 09/495,406

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 876

; TYPE: DNA

; ORGANISM: Campylobacter jejuni

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(876)

; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase

```
; OTHER INFORMATION: Campylobacter sialyltransferase II (estII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus
US-10-961-882-4
```

```
Query Match          96.3%; Score 844; DB 63; Length 876;
Best Local Similarity 97.7%; Pred. No. 8.9e-185;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAGTTTATTTCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
DB |||||||
QY 1 ATGAAAAAGTTTATTTCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
DB |||||||
QY 61 CTACCAAAATGATTTTGTATGATGATTAATCAATTTTATTTTGAAGATAAATCTAT 120
DB |||||||
QY 61 CTACCAAAATGATTTTGTATGATGATTAATCAATTTTATTTTGAAGATAAATCTAT 120
DB |||||||
QY 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATCTAC 180
DB |||||||
QY 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATCTAC 180
DB |||||||
QY 181 ACTTTAAAAATTTAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
DB |||||||
QY 181 ACTTTAAAAATTTAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
DB |||||||
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAATTTTACGATTTATTTTCC 300
DB |||||||
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAATTTTACGATTTATTTTCC 300
DB |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
DB |||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
DB |||||||
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTTATATGTCAGTAGCT 420
DB |||||||
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTTATATGTCAGTAGCT 420
DB |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480
DB |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480
DB |||||||
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
DB |||||||
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540
DB |||||||
QY 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
DB |||||||
QY 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
DB |||||||
QY 601 GAAAAAATTTTAAAGAAATTTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 660
DB |||||||
QY 601 GAAAAAATTTTAAAGAAATTTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 660
DB |||||||
QY 661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 720
DB |||||||
QY 661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTTT 720
DB |||||||
QY 721 AAAGATATCTATACCTTTCTAGTGGGCTTATGGAATTTTCAAAATTTTAAATTTT 780
DB |||||||
QY 721 AAAGATATCTATACCTTTCTAGTGGGCTTATGGAATTTTCAAAATTTTAAATTTT 780
DB |||||||
QY 781 AAAAAATTTAAAGAAATTTTAAAGAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTT 840
DB |||||||
QY 781 AAAAAATTTAAAGAAATTTTAAAGAAATTTTAACTAGTCTTAAAGCTTTTAAAGCTTT 840
DB |||||||
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAATATA 876
DB |||||||
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAATATA 876
DB |||||||
```

## RESULT 63

US-10-962-235-4

; Sequence 4, Application US/10962235



GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,235  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
; OTHER INFORMATION: biosynthesis locus)  
US-10-962-235-4

Query Match 96.3%; Score 844; DB 63; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTCTTTAAAGAAATGATTATTTCAAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTCTTTAAAGAAATGATTATTTCAAGG 60  
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTTAAAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGGAGCAATACTAC 180  
DB 121 CTGTTAAAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGGAGCAATACTAC 180  
QY 181 ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTTAAT 240  
DB 181 ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCTCT 300  
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCTCT 300  
QY 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
DB 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
QY 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTCAAGAAATTTATCTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCT 420  
QY 421 ATAGCCCTTAGATACAAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTTAGATACAAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCATTGCTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCATTGCTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTGCACTATATCGGACATAGTAAATACAGATATATAAAGCTTTTGAATTTCTA 600  
DB 541 GATCGCTGCACTATATCGGACATAGTAAATACAGATATATAAAGCTTTTGAATTTCTA 600

QY 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660  
DB 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720  
QY 721 AAAGATATATCATATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780  
DB 721 AAAGATATATCATATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780  
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATTAAGAGATCTATTAAAGATTA 840  
DB 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATTAAGAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAAATAA 876

## RESULT 64

US-10-962-334-4  
; Sequence 4, Application US/10962334  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,334  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))  
; OTHER INFORMATION: biosynthesis locus)  
US-10-962-334-4

Query Match 96.3%; Score 844; DB 63; Length 876;  
Best Local Similarity 97.7%; Pred. No. 8.9e-185;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTCTTTAAAGAAATGATTATTTCAAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTCTTTAAAGAAATGATTATTTCAAGG 60  
QY 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTTAAAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGGAGCAATACTAC 180  
DB 121 CTGTTAAAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGGAGCAATACTAC 180  
QY 181 ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTTAAT 240  
DB 181 ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATTTGTTCTTAAT 240

181	Db	ACITTTAAACATTTAATCCAAAAATCAAGAAATAGAGACCGAACTAAATTATGTGTTCTAAT	241
241	Qy	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTGAAAAACTTTTTACGATTAATTTCTCT	300
241	Db	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTGAAAAACTTTTTACGATTAATTTCTCT	300
301	Qy	GATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAATTTAAATGCTATATTTTAA	360
301	Db	GATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAATTTAAATGCTATATTTTAA	360
361	Qy	TTTTCACGAAATTTAATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTCAGTAGCC	420
361	Db	TTTTCACGAAATTTAATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTCAGTAGCT	420
421	Qy	ATAGCCCTAGGATCAAAAGAAATTAATCTCTTCGGGAATTGATTTTATCAAAATCGGTCA	480
421	Db	ATAGCCCTAGGATCAAAAGAAATTAATCTCTTCGGGAATTGATTTTATCAAAATCGGTCA	480
481	Qy	TCCTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
481	Db	TCCTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAACTGCGTTCCTGATTTTAAAAAT	540
541	Qy	GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
541	Db	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
601	Qy	GAAAAAACTTACAAAAATAAAACTATATTGCTTATGCTCCTTAATAGTCTTTTACGAATTTT	660
601	Db	GAAAAAACTTACAAAAATAAAACTATATTGCTTATGCTCCTTAATAGTCTTTTACGAATTTT	660
661	Qy	ATGAACTAGCGCCAAATTTTAAATTTCAAATTTTATCATACAAGAAAAATAACTACACT	720
661	Db	ATGAACTAGCGCCAAATTTTAAATTTCAAATTTTATCATACAAGAAAAATAACTACACT	720
721	Qy	AAAGATATATCTATACCTTTCTAGTGAGGCTTATGGAAAAATTTTCAAATAATTAATTTT	780
721	Db	AAAGATATATCTATACCTTTCTAGTGAGGCTTATGGAAAAATTTTCAAATAATTAATTTT	780
781	Qy	AAAAAATAAAAAATAAAGAAAAATGTTTATTAACAAGTTGATAAAAAAGATCTATTAAGATTA	840
781	Db	AAAAAATAAAAAATAAAGAAAAATATTTATTAACAAGTTGATAAAAAAGATCTATTAAGATTA	840
841	Qy	CCTAGTGATATAAAGCATTTATTTCAAAGGAAATAA	876
841	Db	CCTAGTGATATAAAGCATTTATTTCAAAGGAAATAA	876

## RESULT 65

US-10-303-161-1

US 10-303-101-1  
: Sequence 1, Application US/10303161

GENERAL INFORMATION:

APPLICANT: Gilbert, Michel

APPLICANT: Wakarchuk, Warren W.

APPLICANT: National Research Council of Canada

**TITLE OF INVENTION: Campylobacter Glycosyltransferases f**

1. TITLE OF INVENTION: Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-000111US

; CURRENT APPLICATION NUMBER: US/1

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US/09/

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US

;  
PRIOR FILING DATE: 2000-01-18

; NUMBER OF SEQ ID NOS: 49

```

; SOFTWARE: I
; GEO ID NO 1

```

```

; SEQ ID NO 1
: LENGTH: 11474

```

```

; LENGTH: 1
; TYPE: DNA

```

;; TYPE: DNA  
; ; ORGANISM: *Campylobacter jejuni*

ORGANISM  
FEATURE.

OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384

; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-303-161-1

Query Match      95.6%;      Score 837.6;      DB 47;      Length 11474;

Best Local Similarity 97.3%; Pred. No. 5.5e-183;

Matches	852;	Conservative	0;	Mismatches	24;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	-----	--------	----	------	----

Qy	1	ATGAAAAAGTTATTATTGCTGCGAAATGGACCAAGTTTAAAGAAATTTGAATTTTCAAGG	60
Db	6048	ATGAAAAAGTTATTATTGCTGCGAAATGGACCAAGTTTAAAGAAATTTGAATTTTCAAGA	6107
Qy	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTTGAAGATAAATACTAT	120
Db	6108	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTTGAAGATAAATACTAT	6167
Qy	121	CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTGTGAGCAATACTAC	180
Db	6168	CTTGGTAAAAAATGCAAGCAGTATTTTTCAATCCTATTCTTTTGTGAACAACTACTAC	6227
Qy	181	ACTTTAAAAATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTGTCTTAAT	240
Db	6228	ACTTTAAAAATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTGTCTTAAT	6287
Qy	241	TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTGTA AAAACTTTTACGATATTTTTCCCT	300
Db	6288	TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTGTA AAAACTTTTACGATATTTTTCCCT	6347
Qy	301	GATGCTCATTTGGATATGATTTTAAAAACAACCTTAAAGAAATTAATGCTTATTTTAAA	360
Db	6348	GATGCTCATTTGGATATGATTTTAAAAACAACCTTAAAGAAATTAATGCTTATTTTAAA	6407
Qy	361	TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGTGCAAGTACC	420
Db	6408	TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGTGCAAGTACC	6467
Qy	421	ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGGTCA	480
Db	6468	ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGGTCA	6527
Qy	481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
Db	6528	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	6587
Qy	541	GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA	600
Db	6588	GATAATTCACCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA	6647
Qy	601	GAAAAAACTTTACAAAATAAAACTATATGCTTATGCTCTAATAGCTTTTGTAGCAAAATTTT	660
Db	6648	GAAAAAACTTTACAAAATAAAACTATATGCTTATGCTCTAATAGCTTTTGTAGCAAAATTTT	6707
Qy	661	ATAGAACTAGCGGCAAAATTTTAAATTTTAAATTTTATCATACAAGAAAAAATAACTACACT	720
Db	6708	ATAGAACTAGCGGCAAAATTTTAAATTTTAAATTTTATCATACAAGAAAAAATAACTACACT	6767
Qy	721	AAAGATATCTCATACCTCTCTAGTGGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT	780
Db	6768	AAAGATATCTCATACCTCTCTAGTGGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT	6827
Qy	781	AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA	840
Db	6828	AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA	6887
Qy	841	CCTAGTGATATAAGCAATTTTTCAAAGGAAAAATAA	876
Db	6888	CCTAGTGATATAAGCAATTTTTCAAAGGAAAAATAA	6923

RESULT 66

US-10-734-719-1

US 10,734,719  
: Sequence 1, Application US/10734719

GENERAL INFORMATION:

APPLICANT: Gilbert, Michel

;; APPLICANT: Wakarchuk, Warren W.  
;; APPLICANT: National Research Council of Canada  
;; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
;; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
;; FILE REFERENCE: 019633-000111US  
;; CURRENT APPLICATION NUMBER: US/10/734,719  
;; CURRENT FILING DATE: 2003-12-11  
;; PRIOR APPLICATION NUMBER: US 09/816,028  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/118,213  
;; PRIOR FILING DATE: 1999-02-01  
;; PRIOR APPLICATION NUMBER: US 09/495,406  
;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 1  
;; LENGTH: 11474  
;; TYPE: DNA  
;; ORGANISM: Campylobacter jejuni  
;; FEATURE:  
;; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
;; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-734-719-1

Query Match 95.6%; Score 837.6; DB 60; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60  
DB 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGA 6107

QY 61 CTACCAATGATTTTGTATGATGTAATCAATTTATTTTGAAGATAAACTACTAT 120  
DB 6108 CTACCAATGATTTTGTATGATGTAATCAATTTATTTTGAAGATAAACTACTAT 6167

QY 121 CTTCGTAATAATGCAAGCAGTGTTCCTACCCCTAAATTTCTCTTTGGAGCAATACTAC 180  
DB 6168 CTTCGTAATAATGCAAGCAGTGTTCCTACCCCTAAATTTCTCTTTGGAGCAATACTAC 6227

QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTTAAT 240  
DB 6228 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTTAAT 6287

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCTCT 300  
DB 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCTCT 6347

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAAATGCTTATTTTAA 360  
DB 6348 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTTAAATGCTTATTTTAA 6407

QY 361 TTTCACGAAATTTATTTCAAATCAAGAAATTTACTCAGGGGTCTATATGTGTGCGAGTAGCC 420  
DB 6408 TTTCACGAAATTTATTTCAAATCAAGAAATTTACTCAGGGGTCTATATGTGTGCGAGTAGCC 6467

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATCGGTCA 480  
DB 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATCGGTCA 6527

QY 481 TCCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAT 540  
DB 6528 TCCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAT 6587

QY 541 GATCGCTCGCACTATATCGGCATAGTAAATATACAGATATAAAGCTTTTGAATTTCTTA 600  
DB 6588 GATCAATTCACACTATATCGGCATAGTAAATATACAGATATAAAGCTTTTGAATTTCTTA 6647

QY 601 GAAAAAATCTTACAAAATAAAACTATTTGCTTTATGCTTATGCTTATGCTTTTACGAAATTTT 660  
DB 6648 GAAAAAATCTTACAAAATAAAACTATTTGCTTTATGCTTATGCTTATGCTTTTACGAAATTTT 6707

QY 661 ATAGAAGTAGCGCCAAATTTTAAATTTCAATTTTATCATACAGAAAAAATAACTACTAC 720

DB 6708 ATAGAAGTAGCGCCAAATTTAAATTTTATCATACAGAAAAAATAACTACTAC 6767  
QY 721 AAAGATATATCATACCTTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTT 780  
DB 6768 AAAGATATATCATACCTTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTT 6827

QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
DB 6828 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 6887

QY 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAAATAA 876  
DB 6888 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAAATAA 6923

RESULT 67  
US-10-820-536-1  
; Sequence 1, Application US/10820536  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-820-536-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60  
DB 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGA 6107

QY 61 CTACCAATGATTTTGTATGATGTAATCAATTTATTTTGAAGATAAACTACTAT 120  
DB 6108 CTACCAATGATTTTGTATGATGTAATCAATTTATTTTGAAGATAAACTACTAT 6167

QY 121 CTTCGTAATAATGCAAGCAGTGTTCCTACCCCTAAATTTCTTTTGGAGCAATACTAC 180  
DB 6168 CTTCGTAATAATGCAAGCAGTGTTCCTACCCCTAAATTTCTTTTGGAGCAATACTAC 6227

QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTTAAT 240  
DB 6228 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTTAAT 6287

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCTCT 300  
DB 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCTCT 6347

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAAATGCTTATTTTAA 360

Db 6348 GATGCTCAITGGGATATGATTTTCAAACAACCTTAAGATTTTAATGCTTATTTAAA 6407  
Qy 361 TTTTCAGCAAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Db 6408 TTTTCAGCAAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 6467  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
Qy 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 6528 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAAATGGCTCCTAAATTTAAAAAT 6587  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAGATACAGATATAAGCTTTAGAAATTTCTA 600  
Db 6588 GATAAATTCACATATATCGGACATAGTAAAGATACAGATATAAGCTTTAGAAATTTCTA 6647  
Qy 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Db 6648 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 6707  
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 6767  
Qy 721 AAAGATATATCTATCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 6768 AAAGATATATCTATCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827  
Qy 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 6828 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887  
Qy 841 CCTAGTGATATAAGCAATTTTCAAAGGAAAAATAA 876  
Db 6888 CCTAGTGATATAAGCAATTTTCAAAGGAAAAATAA 6923

RESULT 68

US-10-821-573-1  
; Sequence 1, Application US/10821573  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,573  
; PRIOR FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-821-573-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;  
Best Local Similarity 97.8%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60  
Db 6048 ATGAAAAAGTTATTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGA 6107  
Qy 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATATAATACTAT 120  
Db 6108 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATATAATACTAT 6167  
Qy 121 CTTGGTAAAAATGCAAGCAGTCTTTTACACCCCTAATTTCTTTCTTTGAGCAATACTAC 180  
Db 6168 CTTGGTAAAAATGCAAGCAGTCTTTTACAAATCTTCTTTTGTGAAACAAATACTAC 6227  
Qy 181 ACTTTAAAAATTTTAAATCCAAAATCAAGAATATGAGACCGAACTAAATTTATGTGTTCTAAT 240  
Db 6228 ACTTTAAAAATTTTAAATCCAAAATCAAGAATATGAGACCGAACTAAATTTATGTGTTCTAAT 6287  
Qy 241 TACAACCAAGCTATCTAGAAAAATGAAAAATTTGTAAAACTTTTTAGATTTTCTCT 300  
Db 6288 TACAACCAAGCTATCTAGAAAAATGAAAAATTTGTAAAACTTTTTAGATTTTCTCT 6347  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAGAATTTAAATGCTTTATTTTAAA 360  
Db 6348 GATGCTCATTTGGGATATGATTTTCAACAACCTTAAGAATTTAAATGCTTTATTTTAAA 6407  
Qy 361 TTTTCAACGAAATTTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Db 6408 TTTTCAACGAAATTTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 6467  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
Qy 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 6528 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAAATTTGGCTCTCTAAATTTAAAAAT 6587  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAGATACAGATATAAGCTTTAGAAATTTCTA 600  
Db 6588 GATAAATTCACATATATCGGACATAGTAAAGATACAGATATAAGCTTTAGAAATTTCTA 6647  
Qy 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Db 6648 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 6707  
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 6767  
Qy 721 AAAGATATATCTATCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 6768 AAAGATATATCTATCTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827  
Qy 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 6828 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887  
Qy 841 CCTAGTGATATAAGCAATTTTCAAAGGAAAAATAA 876  
Db 6888 CCTAGTGATATAAGCAATTTTCAAAGGAAAAATAA 6923

RESULT 69

US-10-821-604-1  
; Sequence 1, Application US/10821604  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604

```
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-821-604-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;
Best Local Similarity 97.3%; Pred. No. 5.5e-183;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATCGACCAAGCTTTAAAGAAATTTGATTATTCAAGG 60
Db 6048 ATGAAAAAGTTATTATTGCTGGAAATCGACCAAGCTTTAAAGAAATTTGATTATTCAAGA 6107

QY 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATPACTAT 120
Db 6108 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATPACTAT 6167

QY 121 CTGGTAAAAAATGCAAGCGAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATPACTAC 180
Db 6168 CTGGTAAAAAATGCAAGCGAGTATTTTACAATCTCTATTTCTTTTGAACAATPACTAC 6227

QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCGGAACTAATATGTTCTTAAT 240
Db 6228 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCGGAACTAATATGTTCTTAAT 6287

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGAAGAAATTTTACGATTTATTTTCT 300
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGAAGAAATTTTACGATTTATTTTCT 6347

QY 301 GAAAAAAGCTTACAAATAAATACTATTGCTTATGCTCTTAATAGTCTTTTACCAATTTT 660
Db 6647 GAAAAAAGCTTACAAATAAATACTATTGCTTATGCTCTTAATAGTCTTTTACCAATTTT 6707

QY 661 ATAGAACTAGCGGCCAAATTTAAATTTCAATTTTATCATAACAAGAAAAAATACTACACT 720
Db 6708 ATAGAACTAGCGGCCAAATTTAAATTTCAATTTTATCATAACAAGAAAAAATACTACACT 6767

QY 721 AAAGATATCTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
```

```
Db 6768 AAAGATATCTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 6827

QY 781 AAAAAAATAAATAAATAAAGAAATGTTTATTACAGTTGATTAAGAGATCTATTAAAGATTA 840
Db 6828 AAAAAAATAAATAAATAAAGAAATGTTTATTACAGTTGATTAAGAGATCTATTAAAGATTA 6887

QY 841 CCTAGTGATATATAAGCATTATTATTCAAAGGAAAAATAA 876
Db 6888 CCTAGTGATATATAAGCATTATTATTCAAAGGAAAAATAA 6923

RESULT 70
US-10-830-825-1
; Sequence 1, Application US/10830825
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011105
; CURRENT APPLICATION NUMBER: US/10/830,825
; CURRENT FILING DATE: 2004-04-24
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-830-825-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;
Best Local Similarity 97.3%; Pred. No. 5.5e-183;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGCAAGCTTTAAAGAAATTTGATTATTCAAGG 60
Db 6048 ATGAAAAAGTTATTATTGCTGGAAATGCAAGCTTTAAAGAAATTTGATTATTCAAGA 6107

QY 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTCTTTTGAAGATAAATACTAT 120
Db 6108 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167

QY 121 CTGGTAAAAAATGCAAGCGAGTGTTTTACACCCCTAAATTTCTTTTGAAGATAAATACTAC 180
Db 6168 CTGGTAAAAAATGCAAGCGAGTATTTTACAATCTCTATTTCTTTTGAACAATACTAC 6227

QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCGGAACTAATATGTTCTTAAT 240
Db 6228 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCGGAACTAATATGTTCTTAAT 6287

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGAAGAAATTTTACGATTTATTTTCT 300
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGAAGAAATTTTACGATTTATTTTCT 6347

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTTAAATGCTTTATTTAAA 360
Db 6348 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTTAAATGCTTTATTTAAA 6407

QY 361 TTTTACGAAAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420
Db 6408 TTTTACGAAAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 6467
```



```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-845-408-1

Query Match      95.6%; Score 837.6; DB 61; Length 11474;
Best Local Similarity 97.3%; Pred. No. 5.5e-183;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAAAGAAATTTGATTATTCAGAGG 60
DB 6048 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAAAGAAATTTGATTATTCAGAG 6107
QY 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 6108 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
DB 6168 CTTGGTAAAAATGCAAGCAGTGTTTACCAATCTCTATTTCTTTTGAACAATACTAC 6227
QY 181 ACTTTAAACATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTAAT 240
DB 6228 ACTTTAAACATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTAAT 5287
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCT 300
DB 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCT 6347
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTTAAAGCTTTATTTAA 360
DB 6348 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTTAAAGCTTTATTTAA 6407
QY 361 TTTACGAAATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTAAT 420
DB 6408 TTTACGAAATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTAAT 5467
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
DB 6528 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 6587
QY 541 GATCCCTCGCACTATATCGGACATAGTAAATATACAGATATATAAGCTTTTACAATTTCTA 600
DB 6588 GATAATTCACACTATATCGGACATAGTAAATATACAGATATATAAGCTTTTACAATTTCTA 6647
QY 601 GAAAAAACTTACAAAATAAAACTATATGCTTTATGCTTAATAGTCTTTTACCAATTTT 660
DB 6648 GAAAAAACTTACAAAATAAAACTATATGCTTTATGCTTAATAGTCTTTTACCAATTTT 6707
QY 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAAGAAAAAATACTACACT 720
DB 6708 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAAGAAAAAATACTACACT 6767
QY 721 AAAGATATACATACCTCTAGTGGCTTATGGAATTTTCAAAAATATTAATTTT 780
DB 6768 AAAGATATACATACCTCTAGTGGCTTATGGAATTTTCAAAAATATTAATTTT 6827
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
DB 6828 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 8487
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876
```

```
DB 6888 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 6923

RESULT 73
US-10-845-412-1
; Sequence 1, Application US/10845412
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-845-412-1

Query Match      95.6%; Score 837.6; DB 61; Length 11474;
Best Local Similarity 97.3%; Pred. No. 5.5e-183;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAAAGAAATTTGATTATTCAGAGG 60
DB 6048 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAAAGAAATTTGATTATTCAGAG 6107
QY 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 6108 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
DB 6168 CTTGGTAAAAATGCAAGCAGTGTTTACCAATCTCTATTTCTTTTGAACAATACTAC 6227
QY 181 ACTTTAAACATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTTAAAT 240
DB 6228 ACTTTAAACATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTTAAAT 6287
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCT 300
DB 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCT 6347
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTTAAAGAAATTTTAAAGCTTTATTTAA 360
DB 6348 GATGCTCATTTGGGATATGATTTTAAACAACCTTTAAAGAAATTTTAAAGCTTTATTTAA 6407
QY 361 TTTACGAAATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTTAAAT 420
DB 6408 TTTACGAAATTTTAAATCAAGAAATATGACCGAACTAATATGTTCTTAAAT 6467
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
```







; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-847-983-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
QY 1 ATGAAAAAGTTATTATGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
DB 6048 ATGAAAAAGTTATTATGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 6107
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 120
DB 6108 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 6167
QY 121 CTGCTGTAATAAATGCAAGCAGTGTGTTTACACCCCTTAATTTCTTCTTGAGCAATPACTAC 180
DB 6168 CTGCTGTAATAAATGCAAGCAGTGTGTTTACCAATCCTATTTCTTTTGTGAACAATACTAC 6227
QY 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 240
DB 6228 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 6287
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 300
DB 6288 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 6347
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTAAA 360
DB 6348 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTAAA 6407
QY 361 TTTCAAGAAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 420
DB 6408 TTTCAAGAAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 6467
QY 421 ATAGCCCTAGCCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 6588 GATATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6647
QY 601 GAAAAAATTTACAAAATTAAGAAATGTTTATTTCAAGAGTTGATATAAAGATCTATTAGATTA 660
DB 6648 GAAAAAATTTACAAAATTAAGAAATGTTTATTTCAAGAGTTGATATAAAGATCTATTAGATTA 6707
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCACAAGAAAAAATAACTACACT 720
DB 6708 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCACAAGAAAAAATAACTACACT 6767
QY 721 AAAGATATATCTACATCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
DB 6768 AAAGATATATCTACATCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 6827
QY 781 AAAAAAATTAATAAATTAAGAAATGTTTATTTCAAGAGTTGATATAAAGATCTATTAGATTA 840
DB 6828 AAAAAAATTAATAAATTAAGAAATGTTTATTTCAAGAGTTGATATAAAGATCTATTAGATTA 8467
QY 841 CCTAGTGATATAAGCAATTTTCAAGAGGAAATAA 876
DB 6888 CCTAGTGATATAAGCAATTTTCAAGAGGAAATAA 6923
```

RESULT 76

US-10-850-125-1

; Sequence 1, Application US/10850125

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Watren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,125  
; CURRENT FILING DATE: 2004-05-19  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-850-125-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
QY 1 ATGAAAAAGTTATTATGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
DB 6048 ATGAAAAAGTTATTATGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 6107
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 120
DB 6108 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGGAAGATAAATACTAT 6167
QY 121 CTGCTGTAATAAATGCAAGCAGTGTGTTTACACCCCTTAATTTCTTCTTGAGCAATPACTAC 180
DB 6168 CTGCTGTAATAAATGCAAGCAGTGTGTTTACCAATCCTATTTCTTTTGTGAACAATACTAC 6227
QY 181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 240
DB 6228 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 6287
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 300
DB 6288 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 6347
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTAAA 360
DB 6348 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTAAA 6407
QY 361 TTTCAAGAAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 420
DB 6408 TTTCAAGAAATTTAATCCAAATCAAGAAATATGAGCGGAACATAATATGTTGTTCTAAT 6467
QY 421 ATAGCCCTAGCCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 480
DB 6468 ATAGCCCTAGCCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6527
QY 481 TCTTATGCTTTTATGATACCAAGAAATTTTAAACAACCTTTTAAACAATTTTAAAAAT 540
DB 6528 TCTTATGCTTTTATGATACCAAGAAATTTTAAACAACCTTTTAAACAATTTTAAAAAT 6587
QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 6588 GATATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6647
QY 601 GAAAAAATTTACAAAATTAAGAAATGTTTATTTCAAGAGTTGATATAAAGATCTATTAGATTA 660
DB 6648 GAAAAAATTTACAAAATTAAGAAATGTTTATTTCAAGAGTTGATATAAAGATCTATTAGATTA 6707
```

QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT 720  
|||||  
Db 6708 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT 6767  
|||||  
QY 721 AAGATATATCTCATACCTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
|||||  
Db 6768 AAGATATATCTCATACCTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827  
|||||  
QY 781 AAAAAATAAAAATTAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 6828 AAAAAATAAAAATTAAGAAATATTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887  
|||||  
QY 841 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 876  
|||||  
Db 6888 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 6923  
|||||

RESULT 77  
US-10-850-807-1  
; Sequence 1, Application US/10850807  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,807  
; CURRENT FILING DATE: 2004-05-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-850-807-1

Query Match 95.6%; Score 837.6; DB 61; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAGTTTAAAGAAATTTGATTATTTCAAGG 60  
|||||  
Db 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAGTTTAAAGAAATTTGATTATTTCAAGA 6107  
|||||  
QY 61 CTACCAAAATGATTTGATGTTATAGATGTAATCAATTTTATTTGAGATAAATACTAT 120  
|||||  
Db 6108 CTACCAAAATGATTTGATGTTATAGATGTAATCAATTTTATTTGAGATAAATACTAT 6167  
|||||  
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
|||||  
Db 6168 CTTGGTAAAAATGCAAGCAGTGTTTACATCTTATTTCTTTTGGACAAATACTAC 6227  
|||||  
QY 181 ACTTTAAAAATTTAATCAAAAAATCAAGATATGAGACCGAATAATATGTTGTTCTAAT 240  
|||||  
Db 6228 ACTTTAAAAATTTAATCAAAAAATCAAGATATGAGACCGAATAATATGTTGTTCTAAT 6287  
|||||  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTAATTTTCT 300  
|||||  
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTAATTTTCT 6347  
|||||  
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360  
|||||  
Db 6348 GATGCTCATTTGGGATATGATTTTTTCAAACAACTTAAAGATTTTAAATGCTTTATTTTAAA 6407  
|||||

QY 361 TTTTACGAAAAATTTATTTTCAATCAAGAAATTTACCTCAGGGTCTTATATGTGTCAGTAGCC 420  
|||||  
Db 6408 TTTTACGAAAAATTTATTTTCAATCAAGAAATTTACCTCAGGGTCTTATATGTGTCAGTAGCC 6467  
|||||  
QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
|||||  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
|||||  
QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
|||||  
Db 6528 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAAAATTTGGCTCCTTAAATTTTAAAAAT 6587  
|||||  
QY 541 GATCGCTCGCACTATATCGGCACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600  
|||||  
Db 6588 GATAATTCACATATATCGGCACATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 6647  
|||||  
QY 601 GAAAAAACTTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTATAGCAAAATTT 660  
|||||  
Db 6648 GAAAAAACTTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTATAGCAAAATTT 6707  
|||||  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTCAATTTTATCATACAAGAAAAATAACTACACT 720  
|||||  
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTTCAATTTTATCATACAAGAAAAATAACTACACT 6767  
|||||  
QY 721 AAGATATATCTCATACCTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
|||||  
Db 6768 AAGATATATCTCATACCTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827  
|||||  
QY 781 AAAAAATAAAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 6828 AAAAAATAAAAAATTTAAAGAAATATTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887  
|||||  
QY 841 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 876  
|||||  
Db 6888 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 6923  
|||||

Query Match 95.6%; Score 837.6; DB 63; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAGTTTAAAGAAATTTGATTATTTCAAGG 60  
|||||



Db 6828 AAAAAAATAAAATTAAGAAAAATATTTATTACAGTTGATAAAGATCTATTAAAGATTAA 6887  
QY 841 CCTAGTGATATAAGCATTATTTCRAAGGAAAAATAA 876  
Db 6888 CCTAGTGATATAAGCATTATTTCRAAGGAAAAATAA 6923

RESULT 80  
US-10-962-334-1  
; Sequence 1, Application US/10962334  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,334  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-962-334-1

Query Match 95.6%; Score 837.6; DB 63; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 5.5e-183;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 6048 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGA 6107

QY 61 CTACCAAAATGATTTGATGATATTAGATGAATCAATTTTATTGGAAGATAAATACTAT 120  
Db 6108 CTACCAAAATGATTTGATGATATTAGATGAATCAATTTTATTGGAAGATAAATACTAT 6167

QY 121 CTTGGTAAAAATGCAACGAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180  
Db 6168 CTTGGTAAAAATGCAAGGAGTATTTTTACAATCTCTATTCTTTTGAACAAATACTAC 6227

QY 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGACCGCAACTAATTAATGTTCTTAAT 240  
Db 6228 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGACCGCAACTAATTAATGTTCTTAAT 6287

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTTCAGGATTATTTCCT 300  
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTTCAGGATTATTTCCT 6347

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 6348 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 6407

QY 361 TTTTCAGGAATTTATTTCATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCGAGTAGCC 420  
Db 6408 TTTTCAGGAATTTATTTCATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCGAGTAGCC 6467

QY 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527

QY 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Db 6528 TCTTATGCTTTTGATATAAACAATAATCTTTTAAAAATTTGGCTCCTTAATTTTAAAAAT 6587  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTTA 600  
Db 6588 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTTA 6647

QY 601 GAAAAAACTTACAAAAATAAACTATATTGCTTATGTCTCTATAGTCTTTTATAGCAAAATTT 660  
Db 6648 GAAAAAACTTACAAAAATAAACTATATTGCTTATGTCTCTACAGCTTTTATAGCAAAATTT 6707

QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACTACT 720  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACTACT 6767

QY 721 AAGATATATCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAAATTT 780  
Db 6768 AAGATATATCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAAATTT 6827

QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 6828 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887

QY 841 CCTAGTGATATAAGCATTATTTCRAAGGAAAAATAA 876  
Db 6888 CCTAGTGATATAAGCATTATTTCRAAGGAAAAATAA 6923

RESULT 81  
US-10-303-161-2  
; Sequence 2, Application US/10303161  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,161  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATON: (1)..(876)  
; OTHER INFORMATION: Bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-303-161-2

Query Match 95.4%; Score 836; DB 47; Length 876;  
Best Local Similarity 97.1%; Pred. No. 6.4e-183;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATCAAAAAAGCTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATCAAAAAAGCTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60

QY 61 CTACCAAAATGATTTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGATGATTTAGATGAATCAATTTTATTTTGAAGATAAATACTAT 120



QY 841 CCTAGTGATATTAAGCATTATTTCAGAGGAAATAA 876  
|||||  
Db 841 CCTAGTGATATTAAGCATTATTTCAGAGGAAATAA 876  
|||||

## RESULT 83

US-10-820-536-2  
; Sequence 2, Application US/10820536  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-820-536-2

Query Match 95.4%; Score 836; DB 61; Length 876;  
Best Local Similarity 97.1%; Pred. No. 6.4e-183;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
  
QY 1 ATGAAAAAGTTATTATTCCTCGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60  
|||||  
Db 1 ATGAAAAAGTTATTATTCCTCGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60  
|||||  
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
|||||  
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
|||||  
QY 121 CTTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
|||||  
Db 121 CTTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
|||||  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGCACTAATATGTGTTCTAAT 240  
|||||  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGCACTAATATGTGTTCTAAT 240  
|||||  
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATATTTCCT 300  
|||||  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATATTTCCT 300  
|||||  
QY 301 GATGCTCATTTGGGATATCATTTTAAACCACTTTAAAGAAATTAATGCTTATTTTAA 360  
|||||  
Db 301 GATGCTCATTTGGGATATCATTTTAAACCACTTTAAAGAAATTAATGCTTATTTTAA 360  
|||||  
QY 361 TTTCAGGAATTTATTTCAATCAAGAATATACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
|||||  
Db 361 TTTCAGGAATTTATTTCAATCAAGAATATACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
|||||  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
|||||

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATGATTTTATCAAAATGGGTCA 480  
|||||  
QY 481 TCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
|||||  
Db 481 TCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
|||||  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
|||||  
Db 541 GATAATTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
|||||  
QY 601 GAAAAAATTTACAAATAAATACTATATGCTTATGCTCTCTAATAGTCTTTTACGAAATTTT 660  
|||||  
Db 601 GAAAAAATTTACAAATAAATACTATATGCTTATGCTCTCTAATAGTCTTTTACGAAATTTT 660  
|||||  
QY 661 ATAGAACTAGCGCGCAATTTTAAATTTCAATTTTATCATACAGAAATAAATACTACT 720  
|||||  
Db 661 ATAGAACTAGCGCGCAATTTTAAATTTCAATTTTATCATACAGAAATAAATACTACT 720  
|||||  
QY 721 AAAGATATATCTCATACCTCTAGTGAGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780  
|||||  
Db 721 AAAGATATATCTCATACCTCTAGTGAGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780  
|||||  
QY 781 AAAAAAATAAAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 781 AAAAAAATAAAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
QY 841 CCTAGTGATATAAAGCATTATTTCAGAGGAAATAA 876  
|||||  
Db 841 CCTAGTGATATAAAGCATTATTTCAGAGGAAATAA 876  
|||||

## RESULT 84

US-10-821-573-2  
; Sequence 2, Application US/10821573  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,573  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-821-573-2

Query Match 95.4%; Score 836; DB 61; Length 876;  
Best Local Similarity 97.1%; Pred. No. 6.4e-183;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
  
QY 1 ATGAAAAAGTTATTATTCCTCGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60  
|||||

Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGCTTTAAAAAGAAATGATTATTCAAGA 60  
Qy 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTATTGAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTATTGAGATAAATACTAT 120  
Qy 121 CTGTGTAATAAATGCAAAAGCAGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGGCAGTATTTTACAATCCCTATTCTTTTGTGAACAATACTAC 180  
Qy 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGACCGAATTAATATGTTCTAAT 240  
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGACCGAATTAATATGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300  
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAAATGCTTATTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTCAACAACCTTAAAGAAATTTAAATGCTTATTTAAA 360  
Qy 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Db 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGGATATGATTTTCAACAACCTTAAAGAAATTTAAATGCTTATTTAAA 360  
Db 481 TCTTATGCTTTTGGATATGATTTTCAACAACCTTAAAGAAATTTAAATGCTTATTTAAA 360  
Qy 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATTTAAATTTAAATTTTATCAACAAGAAATTTTACGATTAATTTT 660  
Db 601 GAAAAAATTTACAAAATTTAAATTTAAATTTTATCAACAAGAAATTTTACGATTAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCAACAAGAAATTTTAAATTTTAA 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCAACAAGAAATTTTAAATTTTAA 720

RESULT 85

US-10-821-604-2  
; Sequence 2, Application US/10821604  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; PRIORITY FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (estII) from C. jejuni  
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))  
; OTHER INFORMATION: biosynthesis locus)  
US-10-821-604-2

Query Match 95.4%; Score 836; DB 61; Length 876;  
Best Local Similarity 97.1%; Pred. No. 64e-183;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGA 60  
Qy 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Qy 121 CTGTGTAATAAATGCAAAAGCAGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGGCAGTATTTTACAATCCCTATTCTTTTGTGAACAATACTAC 180  
Qy 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGACCGAATTAATATGTTCTAAT 240  
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGACCGAATTAATATGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCT 300  
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAAATGCTTATTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTCAACAACCTTAAAGAAATTTAAATGCTTATTTAAA 360  
Qy 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Db 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGGATATGATTTTAAACAACCTTAAAGAAATTTAAATTTTAA 540  
Db 481 TCTTATGCTTTTGGATATGATTTTAAACAACCTTAAAGAAATTTAAATTTTAA 540  
Qy 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATTTAAATTTAAATTTTATCAACAAGAAATTTTAAATTTTAA 660  
Db 601 GAAAAAATTTACAAAATTTAAATTTAAATTTTATCAACAAGAAATTTTAAATTTTAA 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCAACAAGAAATTTTAAATTTTAA 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCAACAAGAAATTTTAAATTTTAA 720



```
QY 721 AAAGATATATCTCATCTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATATCTCATCTCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAATAAAAATTAAGAAATGTTTATTACAAGTGTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATATTTATTACAAGTGTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 86
US-10-830-825-2
; Sequence 2, Application US/10830825
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/830,825
; CURRENT FILING DATE: 2004-04-24
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-830-825-2

Query Match 95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60
QY 61 CTACCAAAATGATTTTGAATGATGTAATCAATTTTATTGAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGAATGATGTAATCAATTTTATTGAGATAAATACTAT 120
QY 121 CTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180
Db 121 CTGGTAAAAAATGCAAGCAGTGTTTTACAAATCTTATCTTTTGGAGCAATACTAC 180
QY 181 ACTTTAAAAATTAATCAAAAATCAAGATATGAGACCGAATTAATATGTTCTTAAT 240
Db 181 ACTTTAAAAATTAATCAAAAATCAAGATATGAGACCGAATTAATATGTTCTTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAATCTTTTACGATTATTTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAATCTTTTACGATTATTTTCC 300
QY 301 GATGCTCAATTTGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
Db 301 GATGCTCAATTTGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
```

```
QY 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTTA 600
QY 601 GAAAAAATTTACAAAAATAAATACTATATGCTTATGCTTAATAGTCTTTTATGCAAAATTTT 660
Db 601 GAAAAAATTTACAAAAATAAATACTATATGCTTATGCTTAATAGTCTTTTATGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTT 780
QY 781 AAAAAAATAAAAAATTAAGAAATTTTATTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTAAGAAATTTTATTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
```

## RESULT 87

```
US-10-830-997-2
; Sequence 2, Application US/10830997
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/830,997
; CURRENT FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-830-997-2
```



```
Query Match      95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGG 60
   |||||
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGA 60

QY 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120
   |||||
Db 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120

QY 121 CTTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGGAGCAATACTAC 180
   |||||
Db 121 CTTGTGTAATAAATGCAAGCAGTGTTTTACAAATCTTATCTCTTCTTGGAGCAATACTAC 180

QY 181 ACTTTTAAACATTTTAAATCAAAATCAAGAATATGAGCCGAACAAATATATGTGTTCTAAT 240
   |||||
Db 181 ACTTTTAAACATTTTAAATCAAAATCAAGAATATGAGCCGAACAAATATATGTGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATTTTTCCT 300
   |||||
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATTTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAAATCTTTATTTAA 360
   |||||
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAAATCTTTATTTAA 360

QY 361 TTTTACGAAAATTTTAAATCAAAAGAAATTTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
   |||||
Db 361 TTTTACGAAAATTTTAAATCAAAAGAAATTTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAAGAAATTTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCA 480
   |||||
Db 421 ATAGCCCTAGGATACAAAAGAAATTTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
   |||||
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
   |||||
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
   |||||
Db 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
```

RESULT 88

```
US-10-845-408-2
; Sequence 2, Application US/10845408
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
```

```
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-845-408-2
```

```
Query Match      95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGG 60
   |||||
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGGATTATTCAAGA 60

QY 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120
   |||||
Db 61 CTACCAAAATGATTTTGGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120

QY 121 CTTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGGAGCAATACTAC 180
   |||||
Db 121 CTTGTGTAATAAATGCAAGCAGTGTTTTACAAATCTTATCTTCTTCTTGGAGCAATACTAC 180

QY 181 ACTTTTAAACATTTTAAATCAAAATCAAGAATATGAGCCGAACAAATATATGTGTTCTAAT 240
   |||||
Db 181 ACTTTTAAACATTTTAAATCAAAATCAAGAATATGAGCCGAACAAATATATGTGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATTTTTCCT 300
   |||||
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATTTTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAAATCTTTATTTAA 360
   |||||
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAAATCTTTATTTAA 360

QY 361 TTTTACGAAAATTTTAAATCAAAAGAAATTTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
   |||||
Db 361 TTTTACGAAAATTTTAAATCAAAAGAAATTTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAAGAAATTTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCA 480
   |||||
Db 421 ATAGCCCTAGGATACAAAAGAAATTTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
   |||||
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
   |||||
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
   |||||
Db 601 GAAAAAACTTACAAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
```

```
QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTAAGAAAAATTTTATTAACAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 89
US-10-845-412-2
; Sequence 2, Application US/10845412
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; PRIOR FILING DATE: 2004-05-12
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-845-412-2

Query Match 95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred.No. 6.4e-183; Indels 0; Gaps 0;
Matches 851; Conservative 0; Mismatches 25;

QY 1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 60
QY 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTCAATTTCTTTCTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACCAATCTTATTTCTTTTGAACAAATACTAC 180
QY 181 ACTTTAAAAACATTTAATCAAAATCAAGAAATATGAGACCGAACTAATATTGTCTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCAAAATCAAGAAATATGAGACCGAACTAATATTGTCTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTGTAATAAATCTTTTACGATTTTTCCT 300
```

```
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAAAATCTTTTACGATTTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTCAAAACAACTTAAAGAAATTTTAATGCTTATTTAAA 360
QY 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGCTGTGCGAGTAC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTTATATGCTGTGCGAGTAC 420
QY 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAACTTACAAAATAAAACTATATGCTTATGCTTCTTAATAGTCTTTTACCAAAATTTT 660
Db 601 GAAAAAACTTACAAAATAAAACTATATGCTTATGCTTCTTAACAGTCTTTTACCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAAAATAACTACACT 720
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780
QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTAAGAAAAATTTTATTAACAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
```

RESULT 90

```
US-10-846-219-2
; Sequence 2, Application US/10846219
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
```

```
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-846-219-2

Query Match      95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGCGAAATGAGCCAAAGTTTAAAGAAATGATTAATCAAGG 60
DB 1 ATGAAAAAGTTATTATTCTGCGAAATGAGCCAAAGTTTAAAGAAATGATTAATCAAGA 60

QY 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
DB 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120

QY 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180
DB 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATATGTGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATATGTGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTAATTTTCC 300
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTAATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360

QY 361 TTTTCAGGAAATTTATTTCATCAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480
DB 361 TTTTCAGGAAATTTATTTCATCAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTACTAAACAAAATTTTGTAAAAATTTGGCTCTCTAATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGTACTAAACAAAATTTTGTAAAAATTTGGCTCTCTAATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAATTTACAAATTAATTTGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660
DB 601 GAAAAAATTTACAAATTAATTTGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 720

QY 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTAA 780
DB 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTTAA 780

QY 781 AAAAAAATAAATAATTAAGAAATGTTTATTAAGAGTTGATTAAGATTAAGATTAAGATTA 840
DB 781 AAAAAAATAAATAATTAAGAAATGTTTATTAAGAGTTGATTAAGATTAAGATTAAGATTA 840

QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876
```

RESULT 91  
US-10-847-983-2  
; Sequence 2, Application US/10847983  
; GENERAL INFORMATION:

```
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-847-983-2
```

```
Query Match      95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGCGAAATGAGCCAAAGTTTAAAGAAATGATTAATCAAGG 60
DB 1 ATGAAAAAGTTATTATTCTGCGAAATGAGCCAAAGTTTAAAGAAATGATTAATCAAGA 60

QY 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
DB 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120

QY 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180
DB 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATATGTGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATATGTGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTAATTTTCC 300
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTAATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360

QY 361 TTTTCAGGAAATTTATTTCATCAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480
DB 361 TTTTCAGGAAATTTATTTCATCAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTACTAAACAAAATTTTGTAAAAATTTGGCTCTCTAATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGTACTAAACAAAATTTTGTAAAAATTTGGCTCTCTAATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
```



OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))  
OTHER INFORMATION: biosynthesis locus)  
US-10-850-807-2

```
Query Match          95.4%; Score 836; DB 61; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGA 60

QY 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
DB 121 CTGTGTAATAAATGCAAGCGAGTATTTTACAATCCTATCTTTTTTTTGAACATACTAC 180

QY 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAATTAATTTATGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAATTAATTTATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAAGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
DB 241 TACAACCAAGCTCATCTAGAAAAAGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTATTTAAA 360

QY 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAATTTACAAATATAAATCTATGCTTATGCTTATAGTCTTTTAGCAAAATTTT 660
DB 601 GAAAAAATTTACAAATATAAATCTATGCTTATGCTTATAGTCTTTTAGCAAAATTTT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATCACTACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATCACTACT 720

QY 721 AAAGATATACATACCTTTAGTCAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
DB 721 AAAGATATACATACCTTTAGTCAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780

QY 781 AAAAAATATAAATTTAAAGAAATTTTATTTAACAAGTTGATAAAAAAGATCTATTAAGATTA 840
DB 781 AAAAAATATAAATTTAAAGAAATTTTATTTAACAAGTTGATAAAAAAGATCTATTAAGATTA 840

QY 841 CCTAGTGATATAACCATTTTCAACGAAATATA 876
DB 841 CCTAGTGATATAACCATTTTCAACGAAATATA 876
```

RESULT 94  
US-10-961-882-2

Sequence 2, Application US/10961882  
GENERAL INFORMATION:  
APPLICANT: Wakarchuk, Michel  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/961,882  
CURRENT FILING DATE: 2004-10-08  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))  
OTHER INFORMATION: biosynthesis locus)  
US-10-961-882-2

```
Query Match          95.4%; Score 836; DB 63; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-183;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGA 60

QY 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
DB 121 CTGTGTAATAAATGCAAGCGAGTATTTTACAATCTTATCTTTTTTTGAACATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAATTAATTTATGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAATTAATTTATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAAGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
DB 241 TACAACCAAGCTCATCTAGAAAAAGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTATTTAAA 360

QY 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTA 600
```

Db 541 GATAATTACATATATCGGACATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
Qy 601 GAAAAAAGCTTACAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAAGCTTACAAAATAAACTATATGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Qy 721 AAAGATATATCTCATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATAATTAATTTT 780  
Db 721 AAAGATATATCTCATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATAATTAATTTT 780  
Qy 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATT 840  
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATT 840  
Qy 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 95  
US-10-962-235-2  
; Sequence 2, Application US/10962235  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,235  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-962-235-2

Query Match 95.4%; Score 836; DB 63; Length 876;  
Best Local Similarity 97.1%; Pred. No. 6.4e-183;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 60  
Qy 61 CTACCAAAATGATTTGATGATTAGTGAATCAATTTTATTTTGAAGATTAATCTAT 120  
Db 61 CTACCAAAATGATTTGATGATTAGTGAATCAATTTTATTTTGAAGATTAATCTAT 120  
Qy 121 CTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTTCTTGAGCAATCTAC 180  
Db 121 CTGGTAAAAAATGCAAGCAGTGTGTTTACCAATCTCTTTCTTTTGAACAATCTAC 180  
Qy 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGACCGAATAATATTATGTTCTAAT 240

Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAATATGAGACCGGAATAATATTATGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTAATAAACTTTTACGATTATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTAATAAACTTTTACGATTATTTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360  
Qy 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
Db 541 GATAATTCACATATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Qy 721 AAAGATATATCTCATACCTTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATAATTAATTTT 780  
Db 721 AAAGATATATCTCATACCTTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATAATTAATTTT 780  
Qy 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTTCAAGTTGATAAAGATCTATTAAAGATT 840  
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTTCAAGTTGATAAAGATCTATTAAAGATT 840  
Qy 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 96  
US-10-962-334-2  
; Sequence 2, Application US/10962334  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,334  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-962-334-2

Query Match      95.4%; Score 836; DB 63; Length 876;
Best Local Similarity 97.1%; Pred. No. 6.4e-185;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy      1  ATGAAAAAGTATTATTGCTGGAATCGACCAAGCTTTAAAGAAATTTGATTATTCAGG 60
Db      1  ATGAAAAAGTATTATTGCTGGAATCGACCAAGCTTTAAAGAAATTTGATTATTCAGA 60

Qy     61  CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db     61  CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

Qy    121  CTTGTAAAAATGCAAGACGAGTGTTTTACACCCCTAATTTCTCTTTTGAGCAATACTAC 180
Db    121  CTTGTAAAAATGCAAGACGAGTGTTTTACAACTCCTATTTCTTTTGTGAACAATACTAC 180

Qy    181  ACTTTAAACATTTAATCCAAAATCAAGAATATGAGCCGAACCTAATATGTTCTAAT 240
Db    181  ACTTTAAACATTTAATCCAAAATCAAGAATATGAGCCGAACCTAATATGTTCTAAT 240

Qy    241  TACAACCAAGCTCATCTAGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300
Db    241  TACAACCAAGCTCATCTAGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

Qy    301  GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTTAAA 360
Db    301  GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTTAAA 360

Qy    361  TTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCAAGTACC 420
Db    361  TTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCAAGTACC 420

Qy    421  ATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Db    421  ATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

Qy    481  TCTATGCTTTTGTATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db    481  TCTATGCTTTTGTATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy    541  GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db    541  GATAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy    601  GAAAAAATTTACAAAATAAAATATTTGCTTATGCTCTTAATAGTCTTTTACCAATTTT 660
Db    601  GAAAAAATTTACAAAATAAAATATTTGCTTATGCTCTTAATAGTCTTTTACCAATTTT 660

Qy    661  ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720
Db    661  ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720

Qy    721  AAAGATATCTATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
Db    721  AAAGATATCTATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780

Qy    781  AAAAAATTTAAAAATTTTAAAGAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db    781  AAAAAATTTAAAAATTTTAAAGAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840

Qy    841  CCTAGTGATAAAGCATTTTTCAAAGGAAAAATAA 876
Db    841  CCTAGTGATAAAGCATTTTTCAAAGGAAAAATAA 876
```

```

RESULT 97
US-10-303-161-6
; Sequence 6: Application US/10303161
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-303-161-6

Query Match      93.6%; Score 820; DB 47; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy      1  ATCAAAAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db      1  ATCAAAAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60

Qy     61  CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db     61  CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy    121  CTTGTAAAAATGCAAGACGAGTGTTTTACACCCCTAATTTCTTTTGTAGCAATACTAC 180
Db    121  CTTGTAAAAATGCAAGACGAGTGTTTTACAACTCCTATTTCTTTTGAACAAATACTAC 180

Qy    181  ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db    181  ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy    241  TACAACCAAGCTCATCTAGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300
Db    241  TTTAACCAAGCTCATCTAGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

Qy    301  GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTTAAA 360
Db    301  GATGCTCATTTGGGATATGATTTTTCACAACTTTAAAGAAATTTCAATGCTTTTAAA 360

Qy    361  TTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCAAGTACC 420
Db    361  TTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCAAGTACC 420

Qy    421  ATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Db    421  ATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

Qy    481  TCTATGCTTTTGTATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db    481  TCTATGCTTTTGTATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy    541  GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db    541  GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
```



541	GATAAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
601	GAIAAAAACTTACAAAAATAAAACTATATTGCTTATGTCCTAAATAGTCTTTTATGCAAAATTTT	660
601	GAIAAAAACTTACGAAATAAAGCTATATTGTTTATGTCCTAACAGATCTTTTATGCAAAATTTT	660
661	ATAGAACTAGGCCAACATTTTAAATTCAAATTTTATCATACAGAAAGAAAAATAACTACACT	720
661	ATAGAACTAGGCCAACATTTTAAATTCAAATTTTATCATACAGAAAGAAAAATAACTATACT	720
721	AAAGATATACTCATACCTTCTAGTCAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT	780
721	AAAGATATACTCATACCTTCTAGTCAGGCTTATGGAAAAATTTTCAAAAAATATTAAATTTT	780
781	AAAAAATATAAAAATTAAGAAAAATGTTTATTATACAGTTTGATAAAAAGATCTATTAGATTTA	840
781	AAAAAATATAAAAATTAAGAAAAATGTTTATTATACAGTTTGATAAAAAGATCTATTAGATTTA	840
841	CCTAGTGATATAAAGCATTTATTTTCAAGGAAAAATAA	876
841	CCTAGTGATATAAAGCATTTATTTTCAAGGAAAAATAA	876

RESULT 98

```

US-10-734-719-6 .
; Sequence 6, Application US/10734719
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011US
; CURRENT APPLICATION NUMBER: US/10/734,719
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US 09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (csII) from C. jejuni serotype O:41
US-10-734-719-6

```

Query Match	93.6%	Score 820;	DB 60;	Length 876;
Best Local Similarity	96.0%;	Pred. No. 3.3e-179;		
Matches 841;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;
Qy	1	ATGAAAAGAGTTATTATTGCTGGAAATCGACCAAGTTTAAAGAGAAATTTGATTATTCAAGG	60	
Db	1	ATGAAAAGAGTTATTATTGCTGGAAATCGACCAAGTTTAAAGAGAAATTTGATTATTCAAGA	60	
Qy	61	CTACCAATGATTTTGATGTTATTAGATGTGTAATCAATTTTATTTTGAAGATAAAATCTAT	120	
Db	61	CTACCAATGATTTTGATGTTATTAGATGCAATCAATTTTATTTTGAAGATAAAATCTAT	120	
Qy	121	CTTGCTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTCTTTGAGCAAACTACTAC	180	
Db	121	CTTGCTAAAAATGCAAGCAGTGTGTTTACAACTCTAGTCCTTTTTTTTGAACAACTACTAC	180	
Qy	181	ACTTTAAAAACATTTTAACTCCAAAATCAAGAAATATGAGACCGCAACTAATTTATGTGTTCTAAT	240	
Db	181	ACTTTAAAAACATTTTAACTCCAAAATCAAGAAATATGAGACCGCAACTAATCATGTGTTCTAAT	240	

241	Qy	TACAA	CCAGCTCATCTAGAAAAATGAAAAATTTTGTGTAAGAACTTTTACGATATATTTTCT	300
241	Db	TTTAA	CCAGCTCATCTAGAAAAATCAAAAAATTTTGTGTAAGAACTTTTACGATATATTTTCT	300
301	Qy	GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTTAAATGCTTATTTTAAA	360	
301	Db	GATGCTCATTTGGGATATGATTTTTTTTCAACAACCTTAAAGAAATTTCAATGCTTATTTTAAA	360	
361	Qy	TTT	CAGAAAATTTATTTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGTGCGATAGCC	420
361	Db	TTT	CAGAAAATTTATTTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGACACATAGCC	420
421	Qy	ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTGATTTTTATCAAAATGGGTCA	480	
421	Db	ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTGATTTTTATCAAAATGGATCA	480	
481	Qy	TCATTATGCTTTTGATACAAACAAGAAAATCTTTTAAACCTAGGCCCTGATTTTAAAAAT	540	
481	Db	TCATTATGCTTTTGATACCAACAAGAAAATCTTTTAAATTTGGCTCCTAAATTTTAAAAAT	540	
541	Qy	GATGCTTCGCACTATATTCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA	600	
541	Db	GATAAATTCACACTATATTCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA	600	
601	Qy	GAAAAACTTTACAAAAATAAAACTATATTGCTTATGTCCTAATAGTCTTTTACCAAAATTTT	660	
601	Db	GAAAAACTTTACGAAAAATAAGACTATATTGTTTATGTCCTAACAGTCTTTTACCAAAATTTT	660	
661	Qy	ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAGAAAAAAAATAAATACACT	720	
661	Db	ATAGAACTAGCGCCAAATTTAAATTCAAAATTTTATCATACAAGAAAAAAAATAAATATACT	720	
721	Qy	AAAGATATACCTCATACCTTCTAGTGAGGCTTATCGAAAAATTTTCAAAAAATATTAATTTT	780	
721	Db	AAAGATATACCTCATACCTTCTAGTGAGGCTTATCGAAAAATTTTCAAAAAATATTAATTTT	780	
781	Qy	AAAAAATAAAAAATTTAAAGAAAAATGTTTATTACAAGTTGATAAAGAGCTCTATTAAGATT	840	
781	Db	AAAAAATAAAAAATTTAAAGAAAAATTTTATTACAAGTTGATAAAGATCTATTAAAGATT	840	
841	Qy	CCTAGTGATATAAGCAATTTTCAAAGGAAAAATAA	876	
841	Db	CCTAGTGATATAAGCAATTTTCAAAGGAAAAATAA	876	

RESULT 99

US-10-820-536-6  
 ; Sequence 6, Application US/10820536  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/820,536  
 ; CURRENT FILING DATE: 2004-04-07  
 ; PRIOR APPLICATION NUMBER: 10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:



```
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-820-536-6

Query Match      93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTTAAAGAAATGATTATTCAAGG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTTAAAGAAATGATTATTCAAGA 60

Qy 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120

Qy 121 CTGCGTAAAAATGCAAGCAGTGTCTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 CTGCGTAAAAATGCAAGCAGTGTCTTACAACTTAAAGAAATCAATGCTTTTGAACAATACTAC 180

Qy 181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGACCGAACTTAATGTGTCTTAAT 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGACCGAACTTAATGTGTCTTAAT 240

Qy 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 TTTAACCAAGCTCATCTAGAAAAATCAAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 GATGCTCATTTGGGATATGATTTTTCAAACAACTTTTAAAGAAATTTCAATGCTTATTTTAAA 360

Qy 361 TTTTCAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 TTTTCAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCCGCTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 GATCGCTCCGCTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600

Qy 541 GATTAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 GATTAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600

Qy 601 GAAAAAATCTTACAAAAATAAAACTATATTGCTTTATGCTCTTAATAGTCTTTTACGAAATTTT 660
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
601 GAAAAAATCTTACGAAATAAAGCTATATTGTTTATGCTCTTAATAGTCTTTTACGAAATTTT 660

Qy 661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATACTACAT 720
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATACTACAT 720

Qy 721 AAAGATATACATACCTTTAGTCAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
721 AAAGATATACATACCTTTAGTCAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780

Qy 781 AAAAAATAAAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAAAGATCTATTAAAGATTA 840
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
781 AAAAAATAAAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCATTTTTCAAAGGAAAAATAA 876
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
841 CCTAGTGATATAAGCATTTTTCAAAGGAAAAATAA 876

RESULT 100
US-10-821-573-6
```

```
; Sequence 6, Application US/10821573
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-821-573-6
```

```
Query Match      93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTTAAAGAAATGATTATTCAAGG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTTAAAGAAATGATTATTCAAGA 60

Qy 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120

Qy 121 CTGCGTAAAAATGCAAGCAGTGTCTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 CTGCGTAAAAATGCAAGCAGTGTCTTACAACTTAAAGAAATCAATGCTTTTGAACAATACTAC 180

Qy 181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGACCGAACTTAATGTGTCTTAAT 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGACCGAACTTAATGTGTCTTAAT 240

Qy 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 TTTAACCAAGCTCATCTAGAAAAATCAAAAATTTTGTAAAAAATTTTACGATTTATTTCCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 GATGCTCATTTGGGATATGATTTTTCAAACAACTTTTAAAGAAATTTCAATGCTTATTTTAAA 360

Qy 361 TTTTCAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 TTTTCAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCCGCTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 GATCGCTCCGCTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
```

Db 541 GATAATTCACATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAAGCTTACAAAAATAAACTATATGCTTTATGCTCTAAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAAGCTTACAAAAATAAAAGCTATATGTTTATGCTCTAAAGCTTTTAGCAAAATTTT 660  
QY 661 ATAGAAGCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAACACT 720  
Db 661 ATAGAAGCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAACACT 720  
QY 721 AAAGATATATCATACCTTCTAGTGGGCTTATGGAATTTTAAAGAAAAATTTTAAAGAAAAATTAATTTT 780  
Db 721 AAAGATATATCATACCTTCTAGTGGGCTTATGGAATTTTAAAGAAAAATTTTAAAGAAAAATTAATTTT 780  
QY 781 AAAAAAATAAAAAATTAAGAAAAATTTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAAGATTA 840  
Db 781 AAAAAAATAAAAAATTAAGAAAAATTTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAAGATTA 840  
QY 841 CCTAGTGATATATAAGCAATTTTAAAGAAAAATAA 876  
Db 841 CCTAGTGATATATAAGCAATTTTAAAGAAAAATAA 876

## RESULT 101

US-10-821-604-6  
; Sequence 6, Application US/10821604  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; PRIOR FILING DATE: 2004-04-08  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-10-821-604-6

Query Match 93.6%; Score 820; DB 61; Length 876;  
Best Local Similarity 96.0%; Pred. No. 3.3e-179;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTTCAAGA 60  
QY 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACAAATCTAGTCTTTTGTGAAACAATACTAC 180  
QY 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGACGGAAGCAATTAATGTGTTCTAAT 240

Db 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGACGGAAGCAATTAATGTGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAGAACTTTTACGATTATTTTCCCT 300  
Db 241 TTTAAACCAAGCTCATCTAGAAAAATCAAAATTTTGTAAAGAACTTTTACGATTATTTTCCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTCAAACCACTTAAAGAAATTAATGCTTATTTTAAA 360  
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTC 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTC 480  
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGACCTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAAGCTTACAAAAATAAACTATATGCTTATGCTCTAAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAAGCTTACAAAAATAAAAGCTATATGTTTATGCTCTAAAGCTTTTAGCAAAATTTT 660  
QY 661 ATAGAAGCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAAGATTA 720  
Db 661 ATAGAAGCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAAGATTA 720  
QY 721 AAAGATATATCATACCTTCTAGTGGGCTTATGGAATTTTAAAGAAAAATTAATTTTAAA 780  
Db 721 AAAGATATATCATACCTTCTAGTGGGCTTATGGAATTTTAAAGAAAAATTAATTTTAAA 780  
QY 781 AAAAAAATAAAAAATTAAGAAAAATTTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAAGATTA 840  
Db 781 AAAAAAATAAAAAATTAAGAAAAATTTTAAATTTTAAATTTTATCATACAAAGAAAAATAAATTAAGATTA 840  
QY 841 CCTAGTGATATATAAGCAATTTTAAAGAAAAATAA 876  
Db 841 CCTAGTGATATATAAGCAATTTTAAAGAAAAATAA 876

## RESULT 102

US-10-830-825-6  
; Sequence 6, Application US/10830825  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Michel  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/830,825  
; CURRENT FILING DATE: 2004-04-24  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-830-825-6

Query Match      93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATCGAACCAAGTTTAAAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATCGAACCAAGTTTAAAAAGAAATGATTATTCAAGA 60

QY 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

QY 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGCCGAACTAATTTGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGCCGAACTAATTTGTGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 300
Db 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACATTTTAAAGAAATTTAATGCTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACATTTTAAAGAAATTTAATGCTTATTTAAA 360

QY 361 TTTTACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 420
Db 361 TTTTACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATATCCAAACCAAGAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATATCCAAACCAAGAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGACACTATATCGGACATAGTAAATATGATATGATATGATATGATATGATATG 600
Db 541 GATCGCTCGACACTATATCGGACATAGTAAATATGATATGATATGATATGATATGATATG 600

QY 601 GAAAAAATTTACAAATATAATTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 660
Db 601 GAAAAAATTTACAAATATAATTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 660

QY 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAAATCACTAC 720
Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAAATCACTAC 720

QY 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780
Db 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT 780

QY 781 AAAAAAATATAAATTTAAAGAAATATTTTATTAAGATTTGATATTAAGATTTA 840
Db 781 AAAAAAATATAAATTTAAAGAAATATTTTATTAAGATTTGATATTAAGATTTA 840

QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAATAA 876
```

RESULT 103

US-10-830-997-6

```
; Sequence 6, Application US/10830997
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/830,997
; CURRENT FILING DATE: 2004-04-22
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-830-997-6
```

```
Query Match      93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATCGAACCAAGTTTAAAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATCGAACCAAGTTTAAAAAGAAATGATTATTCAAGA 60

QY 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

QY 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db 121 CTGTTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGCCGAACTAATTTGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAATCCAAATCAAGAAATATGAGCCGAACTAATTTGTGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 300
Db 241 TTTTACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACATTTTAAAGAAATTTAATGCTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACATTTTAAAGAAATTTAATGCTTATTTAAA 360

QY 361 TTTTACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 420
Db 361 TTTTACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATATCCAAACCAAGAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATATCCAAACCAAGAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGACACTATATCGGACATAGTAAATATGATATGATATGATATGATATGATATG 600
Db 541 GATCGCTCGACACTATATCGGACATAGTAAATATGATATGATATGATATGATATGATATG 600
```

```
Db 541 GATAATTCACCTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAAATATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAAATATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAAATACTACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAAATACTACT 720
Qy 721 AAAGATATATCTATACCTTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780
Db 721 AAAGATATATCTATACCTTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780
Qy 781 AAAAAAATAAAATTAAGAAATGTTTATTAACAAGTTGATTAAGATCTATTAAGATTA 840
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTAACAAGTTGATTAAGATCTATTAAGATTA 840
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876

RESULT 104
US-10-845-408-6
; Sequence 6, Application US/10845408
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstii) from C. jejuni serotype O:41
US-10-845-408-6

Query Match 93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAAG 60
Qy 61 CTACCAATGATTTTATGATGATTTATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTATGATGATTTATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGGAGCAATACTAC 180
Db 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGGAGCAATACTAC 180
Qy 181 ACTTTAAAAACATTTTAAATCAAGAAATGAGACCGCAACTAATATGTTGTTCTAAT 240
Db 181 ACTTTAAAAACATTTTAAATCAAGAAATGAGACCGCAACTAATATGTTGTTCTAAT 240
```

```
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTCT 300
Db 241 TTTTAAACCAAGCTCATCTAGAAAAATCAAAAATTTTGTAAAAAATTTTACGATTTATTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTTAAA 360
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGACAGTACCC 420
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGACAGTACCC 420
Qy 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGCATACCAAAAGAAATTTCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGCATACCAAAAGAAATTTCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACCTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTTCTA 600
Db 541 GATAATTTACACCTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAAATTTATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAAATTTATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTACT 720
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTACT 720
Qy 721 AAAGATATATCTATACCTTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780
Db 721 AAAGATATATCTATACCTTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780
Qy 781 AAAAAAATAAAATTAAGAAATGTTTATTAACAAGTTGATTAAGATCTATTAAGATTA 840
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTAACAAGTTGATTAAGATCTATTAAGATTA 840
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876

RESULT 105
US-10-845-412-6
; Sequence 6, Application US/10845412
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
```



QY 601 GAAAAAAGCTTCAAAATAAAAGCTATATGCTTATGCTCTTAATGCTTTTAGCAAAATTTT 660  
DB 601 GAAAAAAGCTTCAAAATAAAAGCTATATGCTTATGCTCTTAATGCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATAACTACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATAACTACACT 720  
QY 721 AAAGATATATCATACACTTCTAGTCAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
DB 721 AAAGATATATCATACACTTCTAGTCAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
QY 781 AAAAAAATAAAATTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
DB 781 AAAAAAATAAAATTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

## RESULT 107

US-10-847-983-6  
; Sequence 6, Application US/10847983  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/847,983  
; PRIOR FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-10-847-983-6

Query Match 93.6%; Score 820; DB 61; Length 876;  
Best Local Similarity 96.0%; Pred. No. 3.3e-179;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTTCAAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTTCAAGA 60  
QY 61 CTACCAATGATTTTATGATGATTTATGATGATTAATCAATTTTATTTGAAATTAATTAAT 120  
DB 61 CTACCAATGATTTTATGATGATTTATGATGATTAATCAATTTTATTTGAAATTAATTAAT 120  
QY 121 CTGTTGAAAAATGCAAGCAAGTTTATACCCCTAAATTTCTTTTTCAGCAATTAAT 180  
DB 121 CTGTTGAAAAATGCAAGCAAGTTTATTAATCTCTAGTCTTTTTCAGCAATTAAT 180  
QY 181 ACTTTAAAAACATTTAATCAAAATCAAGATATGAGCCGAACCTAATATGTTCTAAT 240  
DB 181 ACTTTAAAAACATTTAATCAAAATCAAGATATGAGCCGAACCTAATATGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTTCTCCT 300

DB 241 TTTAACCAAGCTCATCTAGAAAAATCAAAATTTTGTAAAAAATTTTACGATTTATTTCT 300  
QY 301 GATGCTCATTTGCGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
DB 301 GATGCTCATTTGCGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
QY 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGTGCAAGTACC 420  
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGTGCAAGTACC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTATACCAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTATACCAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATTAATTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAAGCTTACAAATAAAAGCTATATGCTTATGCTCTTAATGCTTTTAGCAAAATTTT 660  
DB 601 GAAAAAAGCTTACAAATAAAAGCTATATGCTTATGCTCTTAATGCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAACTACTACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAACTACTACT 720  
QY 721 AAAGATATATCATACACTTCTAGTCAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
DB 721 AAAGATATATCATACACTTCTAGTCAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
QY 781 AAAAAAATAAAATTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
DB 781 AAAAAAATAAAATTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

## RESULT 108

US-10-850-125-6  
; Sequence 6, Application US/10850125  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,125  
; CURRENT FILING DATE: 2004-05-19  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41

US-10-850-125-6

```

Query Match      93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAGAAATTTGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAGAAATTTGATTATTCAAGA 60

QY 61 CTACCAATGATTTTGGATGTTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTTGGATGTTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCGAGTGTTTTACACCCCTAATTTCTTCTTTTGAGCAATACTAC 180
DB 121 CTTGGTAAAAATGCAAGCGAGTATTTTACAACTCTAGTCTTTTGTGAACAATACTAC 180

QY 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGAGCCGAACCTAATTTATGTTCTAAT 240
DB 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGAGCCGAACCTAATCATGTGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCC 300
DB 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCC 300

QY 301 GATGCTCAATTTGGGATGATTTTAAAACAACTTTAAAGAAATTTAAATGCTTTTAA 360
DB 301 GATGCTCAATTTGGGATGATTTTAAAACAACTTTAAAGAAATTTAAATGCTTTTAA 360

QY 361 TTTTCAAGAAATTTTCAATCAAGAAATTTTACCGGGTCTATATGTCAGTAGGCC 420
DB 361 TTTTCAAGAAATTTTCAATCAAGAAATTTTACCGGGTCTATATGTCAGTAGGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATCGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATCGGTCA 480

QY 481 TCTTATGCTTTTGATACAAATCAAGAAATCTTTTAAACCTAGCCCTTGATTTTAA 540
DB 481 TCTTATGCTTTTGATACAAATCAAGAAATCTTTTAAACCTAGCCCTTGATTTTAA 540

QY 541 GATGCTCGACTATATCGGCATAGTAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATGCTCGACTATATCGGCATAGTAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTTAAAGAAATTTTAAATTTTATATCAAGAAAAATTTTAAAGCTTTTAA 660
DB 601 GAAAAAATTTTAAAGAAATTTTAAATTTTATATCAAGAAAAATTTTAAAGCTTTTAA 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATATCAAGAAAAATTTTAAAGCTTT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATATCAAGAAAAATTTTAAAGCTTT 720

QY 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAGAAAAATTTTAAATTTT 780
DB 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAGAAAAATTTTAAATTTT 780

QY 781 AAAAATAAATAAATAAAGAAATTTTATTTATACAGTTGATAAAGATCTATTAGATTA 840
DB 781 AAAAATAAATAAATAAAGAAATTTTATTTATACAGTTGATAAAGATCTATTAGATTA 840

QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATTA 876
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATTA 876
```

RESULT 109

```

US-10-850-807-6
; Sequence 6, Application US/10850807
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
```

```

; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstii) from C. jejuni serotype O:41
US-10-850-807-6
```

```

Query Match      93.6%; Score 820; DB 61; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAGAAATTTGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAGAAATTTGATTATTCAAGA 60

QY 61 CTACCAATGATTTTGGATGTTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTTGGATGTTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCGAGTGTTTTACACCCCTAATTTCTTCTTTTGAGCAATACTAC 180
DB 121 CTTGGTAAAAATGCAAGCGAGTATTTTACAACTCTAGTCTTTTGTGAACAATACTAC 180

QY 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGAGCCGAACCTAATTTATGTTCTAAT 240
DB 181 ACTTTAAACATTTTAATCCAAATCAAGAATATGAGCCGAACCTAATCATGTGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCC 300
DB 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCC 300

QY 301 GATGCTCAATTTGGGATGATTTTAAAACAACTTTAAAGAAATTTAAATGCTTTTAA 360
DB 301 GATGCTCAATTTGGGATGATTTTAAAACAACTTTAAAGAAATTTAAATGCTTTTAA 360

QY 361 TTTTCAAGAAATTTTCAATCAAGAAATTTTACCGGGTCTATATGTCAGTAGGCC 420
DB 361 TTTTCAAGAAATTTTCAATCAAGAAATTTTACCGGGTCTATATGTCAGTAGGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATCGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATCGGTCA 480

QY 481 TCTTATGCTTTTGATACAAATCAAGAAATCTTTTAAACCTAGCCCTTGATTTTAA 540
DB 481 TCTTATGCTTTTGATACAAATCAAGAAATCTTTTAAACCTAGCCCTTGATTTTAA 540

QY 541 GATGCTCGACTATATCGGCATAGTAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATGCTCGACTATATCGGCATAGTAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTTAAAGAAATTTTAAATTTTATATCAAGAAAAATTTTAAAGCTTTTAA 660
DB 601 GAAAAAATTTTAAAGAAATTTTAAATTTTATATCAAGAAAAATTTTAAAGCTTTTAA 660
```



QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAAATACTACACT 720  
|||||  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAAATACTACACT 720  
|||||  
QY 721 AAAGATATACCTCATCTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
|||||  
Db 721 AAAGATATACCTCATCTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
|||||  
QY 781 AAAAAATAAAAATTTAAAGAAATCTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 781 AAAAAATAAAAATTTAAAGAAATCTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
QY 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876  
|||||  
Db 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876  
|||||

## RESULT 110

US-10-961-882-6  
; Sequence 6, Application US/10961882  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/961,882  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-10-961-882-6

Query Match 93.6%; Score 820; DB 63; Length 876;  
Best Local Similarity 96.0%; Pred. No. 3.3e-179;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCGCGAAATGGACCAAGTTTAAAGAAATTCGATTATTCACGG 60  
Db 1 ATGAAAAAGTTATTATTCGCGAAATGGACCAAGTTTAAAGAAATTCGATTATTCACGA 60  
QY 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
QY 121 CTGTTGTAATAAATGCAAGCAGTGTGTTACACCCCTAAATTCCTCTTGTGAGCAATAC 180  
Db 121 CTGTTGTAATAAATGCAAGCAGTGTGTTACACCCCTAAATTCCTCTTGTGAGCAATAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAGATATGACCGCACTAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAGATATGACCGCACTAATATGTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300  
Db 241 TTTACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTA 360

Db 301 GATGCTCATTTGGGATATGATTTTTTCAACCAACTTAAAGAAATTTCAATGCTTTATTTAAA 360  
QY 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420  
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGTCAGTAGCC 420  
QY 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATGATTTTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATGATTTTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTTAAAAAT 540  
QY 541 GATCGCTCGCAGTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAAATTTCTTA 600  
Db 541 GATAATTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAAATTTCTTA 600  
QY 601 GAAAAAACTTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTACGAAATTTT 660  
Db 601 GAAAAAACTTTACGAAATTAAGCTATATGTTTATGCTCTTAACAGTCTTTTACGAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAAGAAAAAATACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAAGAAAAAATACTATACT 720  
QY 721 AAAGATATACCTCATCTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 721 AAAGATATACCTCATCTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAATAAAAATTTAAAGAAATCTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAAAATTTAAAGAAATCTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAAGCAATTTTCAAGGAAAAATAA 876  
|||||

## RESULT 111

US-10-962-235-6  
; Sequence 6, Application US/10962235  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/962,235  
; CURRENT FILING DATE: 2004-10-08  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-10-962-235-6

Query Match 93.6%; Score 820; DB 63; Length 876;  
Best Local Similarity 96.0%; Pred. No. 3.3e-179;



```
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGCTTTAAAGAAATGATTATTCAAGG 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGCTTTAAAGAAATGATTATTCAAGA 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTAAAGATAAATACTAT 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTAAAGATAAATACTAT 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 CTGTTGTAATAATGCAAGAGTGTGTTTACCCCTAAATTTCTTTGAGCAATACATAC 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 CTGTTGTAATAATGCAAGAGTGTGTTTACCAATCTAGTCTTTTGAACAATACTAC 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATGAGCGAACTAATGTTCTAAT 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATGAGCGAACTAATGTTCTAAT 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAGAACTTTTACGATATTCTCT 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTGTAAGAACTTTTACGATATTCTCT 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATCAATGCTTTTAA 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATCAATGCTTTTAA 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 TTTACGAAATTTTAAATCAAGAAATTTACCTCAGGGTCTATGTTGCAATAGCC 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 TTTACGAAATTTTAAATCAAGAAATTTACCTCAGGGTCTATGTTGCAATAGCC 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGTCA 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGTCA 480
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAA 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAA 540
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 GATCGCTGCACATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 GATAATTCACATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 601 GAAAAAACTTACAAAATAAAAGCTATATGTTTATGCTTAAAGCTCTTTTAGCAAAATTTT 660
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 ATAGAACTAGCGCCCAAAATTTAAATTTAAATTTTATCATACAGAAAAATAAATACTAC 720
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 ATAGAACTAGCGCCCAAAATTTAAATTTAAATTTTATCATACAGAAAAATAAATACTAC 720
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 112

```
US-10-962-334-6
; Sequence 6, Application US/10962334
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
```

```
; CURRENT APPLICATION NUMBER: US/10/962,334
; CURRENT FILING DATE: 2004-10-08
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-962-334-6
```

```
Query Match 93.6%; Score 820; DB 63; Length 876;
Best Local Similarity 96.0%; Pred. No. 3.3e-179;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGCTTTAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGTATTATTGCTGGAATGACCAAGCTTTAAAGAAATGATTATTCAAGA 60
QY 61 CTACCAATGATTTTATGATGTAATTTAGATGTAATCAATTTTAAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTATGATGTAATTTAGATGTAATCAATTTTAAAGATAAATACTAT 120
QY 121 CTGTTGTAATAATGCAAGAGTGTGTTTACCCCTAAATTTCTTTGAGCAATACATAC 180
Db 121 CTGTTGTAATAATGCAAGAGTGTGTTTACCAATCTAGTCTTTTGAACAATACTAC 180
QY 181 ACTTTAAACATTTTAAATCAAGAAATTTAGACCGAACTAATGTTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCAAGAAATTTAGACCGAACTAATGTTGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTCTCT 300
Db 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTCTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTCAATGCTTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTCAATGCTTTTAA 360
QY 361 TTTACGAAATTTTAAATCAAGAAATTTACCTCAGGGTCTATGTTGCAATAGCC 420
Db 361 TTTACGAAATTTTAAATCAAGAAATTTACCTCAGGGTCTATGTTGCAATAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAA 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAA 540
QY 541 GATCGCTGCACATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATAATTCACATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAACTTACAAAATAAAAGCTATATGTTTATGCTTAAAGCTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCCAAAATTTAAATTTAAATTTTATCATACAGAAAAATAAATACTAC 720
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTAAATTTTATCATACAGAAAAATAAATACTAC 720
```

Qy	721	AAAGATATACTACATACCTTCTAGTGAGGCTTATCGAAAAATTTTCAAAAAATATTAAATTTT	780
Db	721	AAAGATATACTACATACCTTCTAGTGAGGCTTATCGAAAAATTTTCAAAAAATATTAAATTTT	780
Qy	781	AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTAA	840
Db	781	AAAAAATAAAAAATTAAGAAAAATATTATTACAAGTTGATAAAAGATCTATTAAAGATTAA	840
Qy	841	CCTAGTGATATAAAGCATTTATTTTCAAAGGAAAAATAA	876
Db	841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAAATAA	876
RESULT 113			
US-10-799-016-1			
; Sequence 1, Application US/10799016			
; GENERAL INFORMATION:			
; APPLICANT: Gilbert, Michel			
; APPLICANT: National Research Council of Canada			
; TITLE OF INVENTION: Lipopolysaccharide alpha-2,3 Sialyltransferase of			
; TITLE OF INVENTION: Campylobacter jejuni and Its Uses			
; FILE REFERENCE: 014137-013210US			
; CURRENT APPLICATION NUMBER: US/10/799,016			
; CURRENT FILING DATE: 2004-03-11			
; PRIOR APPLICATION NUMBER: US/10/058,636			
; PRIOR FILING DATE: 2002-01-29			
; PRIOR APPLICATION NUMBER: US/09/272,960			
; PRIOR FILING DATE: 1999-03-18			
; PRIOR APPLICATION NUMBER: US 60/078,891			
; PRIOR FILING DATE: 1998-03-20			
; PRIOR APPLICATION NUMBER: US 09/272,960			
; PRIOR FILING DATE: 1999-03-18			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 1293			
; TYPE: DNA			
; ORGANISM: Campylobacter jejuni			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(1293)			
; OTHER INFORMATION: Campylobacter jejuni OH4384 cst-I gene			
; OTHER INFORMATION: alpha2,3-sialyltransferase			
US-10-799-016-1			
Query Match 35.6%; Score 311.6; DB 60; Length 1293;			
Best Local Similarity 61.1%; Pred. No. 2.5e-61;			
Matches 522; Conservative 0; Mismatches 329; Indels 3; Gaps 1			
Qy	1	ATGAAAAAGTTATTATTCCTCGGAAATGGACCAAGTTTAAAGAATTTGATTATTCAAGG	60
Db	46	ATGCAAAATAATAATCATAGCAGGAAATGGACCTAGCCCTAAAAAATATTAAATTATAAAGA	105
Qy	61	CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTCAAGATAAAATACTAT	120
Db	106	CTGCTAGAGATATGATGTTTTTAGGTGTACACAGTTTTATTTTGAAGATAAGATTATTAT	165
Qy	121	CTTGGTAAAAAATGCAAGCAGTGTFTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC	180
Db	166	TTAGCAAAAAAGATTAAGACGAGTATTTTTTAATCCTGGTGCTTTTATCAACAGATATCAC	225
Qy	181	ACTTTAAACATTTAATCCAAAATCAAGAATATGAGCCGACATAATTATGTGTTCTAAT	240
Db	226	ACTGCAAAACCACTTATACATAAAAAATGAGTATGAAATATAAAAAATTTTTTTCGCTCTCA	285
Qy	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGAAAAAATTTTTTACGATTATTTTCCT	300
Db	286	TTTAAATTTACCTTTTATTGAAGCAATGATTTTTTTTACATCAATTTTATTAATTTTTTCCC	345
Qy	301	GATGCTCATTTGGGATATGATTTTTTTTAAACCAACTTAAGAATTTTAATGCTTTATTTTAA	360
Db	346	GATSCAAAACCTTGGCTATGAGTTATTTCGAAACCTTAAAGAATTTTATGCTTTATATAAA	405

REFERENCE/DOCKET NUMBER: 1488.0140000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-426-787-1

Query Match 32.7%; Score 286.6; DB 13; Length 1830121;  
Best Local Similarity 60.6%; Pred. No. 1.2e-54;  
Matches 529; Conservative 0; Mismatches 329; Indels 15; Gaps 3;  
QY 10 GTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTAATCAAGGCTACCAAT 69  
Db 379506 GTCATTATGTCAGGTAATGGAACAAGTTTAAATCAATTGACTATAGTTTATTACCTAAA 379447  
QY 70 GATTTCGATGTTTAGATGTAATCAATTTTATTTTGAAGATAAATACATCTTGGTAAA 129  
Db 379446 GATTATGATGTTTCCGTTGCAATCAATTTTATTTTGAAGATCAATTTTCTTGGCAAG 379387  
QY 130 AAATGCAAGCAGTGTCTTTACACCCCTAAATTTCTTTTGAGCAATCTACACTTTAAA 189  
Db 379386 AAAATAAAAAAGTATTTTAAATGTTCTACAATTTTGGCAATCTATACCTTTATG 379327  
QY 190 CATTTAATCCAAAATCAAGAAATATGAGACCGAACTTAATTTGTTGTTCTAAATACACCAA 249  
Db 379326 CAATTAATAAAAAATGAATATGATGCTGATATTTATTTATCATCTTTTGTCAAT 379267  
QY 250 GCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGTATTTTCCGTGATCTCAT 309  
Db 379266 TTAGGAGATTTCAGAAATTAAGAAAATTAAGAAATGTACAAAATTTACTAACACAAGTTGAT 379207  
QY 310 TTGGGATATGATTTTAAACAACTTAAGAAATTTAATGCTTATTTAAATTTTCAACAA 369  
Db 379206 ATTGGACATTTATTTAAACAAGTACCCGCTTTGATGCTTATTACAATATAACGAA 379147  
QY 370 ATTTTATTCATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCCATAGCCCTA 429  
Db 379146 TTATATGAATATAGAGAAATTTACATCAGCGTTTATATGTGTCAGTGGCACTGTAATG 379087  
QY 430 GGATACAAAGAAATTTATCTTTTCGGGAATGATTTTATC---AAAATGGGTATCTTAT 486  
Db 379086 GGTATTAAGATCTTTATTTAACAGTATTGATTTTATCAAGAAAAGGGAATCCTTAC 379027  
QY 487 GCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAATATGATCGC 546  
Db 379026 GCATTTTCATCATCAAAAAGAAATATTTAATTAATTTATACCTTCTTTTTCACAAAATAA 378967  
QY 547 TCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTAGAAAA 606  
Db 378966 AGTCAAGCGATATCCATCTTATGGAATATGATTTAATGACATTTATTTTACAAAA 378907  
QY 607 ACTTACAAAAATAAACTATTTGCTTATGCTTAATAGTCTTTTACGAAATTTTATAGAA 666  
Db 378906 CATTATGAGTAAATATTTATTTGCAATTCGCAAGAAAGTCTCTATGTAATTTTTCCT 378847  
QY 667 CTAGGGCGAAATTTAAATTC---AATTTTATCATCAAGAAAATAAAATACACTACTAAA 723  
Db 378846 TTATCACCACCTGAATAACCCAATTTACTTTTATCTCGAAGAAAAGAAAATTTACACAA 378787  
QY 724 GATATACTCATACCTCTCTAGTGAGCTTATGAAAAATTTTCAAAAAATTAATTTTAAA 783  
Db 378786 GATATTTTAAATCC-----GCCGAAGTTTGTATATAAAAAATTTGGTATATATTC 378736  
QY 784 AAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTTATTAAGATTTACCT 843  
Db 378735 AAACCAAGAAATTTACCAAAATCTGATTTTTCGGTTGATCTGGATATATTTACGTTTACCT 378676  
QY 844 AGTGATATAAGCATATTTTCAAGGAAAAATAA 876

Db 378675 AATGATATAAAACACGCCCTTAAATCAAGAAAA 378643  
RESULT 115  
US-10-329-670-1/c  
; Sequence 1, Application US/10329670  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra  
; FILE REFERENCE: PB186P1  
; CURRENT APPLICATION NUMBER: US/10/329,670  
; CURRENT FILING DATE: 2002-12-24  
; PRIOR APPLICATION NUMBER: US 09/643,990  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: US 08/487,429  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: US 08/426,787  
; PRIOR FILING DATE: 1995-04-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1830121  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (4747)..(4747)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (9921)..(9921)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (10150)..(10150)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (29298)..(29298)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (36543)..(36543)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (36551)..(36551)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (36636)..(36636)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (40808)..(40810)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (44416)..(44416)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (44905)..(44905)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (44975)..(44975)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (45593)..(45593)

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:09:43 ; Search time 999.616 Seconds  
(without alignments)  
10768.108 Million cell updates/sec

Title: US-10-734-719-8

Perfect score: 876

Sequence: 1 atgaaaaagttatttgc.....attatttcaaggaataaa 876

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq:\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11A\_PUBCOMB.seq:\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11B\_PUBCOMB.seq:\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11C\_PUBCOMB.seq:\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US11D\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	876	100.0	876	3	US-09-816-028A-8
2	876	100.0	876	7	US-10-830-161-8
3	876	100.0	876	7	US-10-303-118-8
4	876	100.0	876	7	US-10-303-128-8
5	876	100.0	876	7	US-10-303-134-8
6	876	100.0	876	7	US-10-303-162-8
7	876	100.0	876	9	US-10-735-419-8
8	876	100.0	876	9	US-10-820-536-8
9	876	100.0	876	9	US-10-845-408-8
10	876	100.0	876	9	US-10-845-412-8
11	876	100.0	876	9	US-10-846-219-8
12	876	100.0	876	9	US-10-821-604-8
13	876	100.0	876	9	US-10-847-983-8
14	876	100.0	876	9	US-10-821-573-8
15	876	100.0	876	9	US-10-850-807-8
16	876	100.0	876	9	US-10-850-125-8
17	876	100.0	876	9	US-10-830-825-8

18	876	100.0	876	10	US-10-962-334-8	Sequence 8, Appli
19	876	100.0	876	10	US-10-830-997-8	Sequence 8, Appli
20	876	100.0	876	10	US-10-962-235-8	Sequence 8, Appli
21	876	100.0	876	10	US-10-961-882-8	Sequence 8, Appli
22	869.8	99.3	873	3	US-09-816-028A-13	Sequence 13, Appli
23	869.8	99.3	873	7	US-10-303-161-13	Sequence 13, Appli
24	869.8	99.3	873	7	US-10-303-118-13	Sequence 13, Appli
25	869.8	99.3	873	7	US-10-303-128-13	Sequence 13, Appli
26	869.8	99.3	873	7	US-10-303-134-13	Sequence 13, Appli
27	869.8	99.3	873	7	US-10-303-162-13	Sequence 13, Appli
28	869.8	99.3	873	9	US-10-735-419-13	Sequence 13, Appli
29	869.8	99.3	873	9	US-10-820-536-13	Sequence 13, Appli
30	869.8	99.3	873	9	US-10-845-408-13	Sequence 13, Appli
31	869.8	99.3	873	9	US-10-846-219-13	Sequence 13, Appli
32	869.8	99.3	873	9	US-10-821-604-13	Sequence 13, Appli
33	869.8	99.3	873	9	US-10-847-983-13	Sequence 13, Appli
34	869.8	99.3	873	9	US-10-821-573-13	Sequence 13, Appli
35	869.8	99.3	873	9	US-10-850-807-13	Sequence 13, Appli
36	869.8	99.3	873	9	US-10-850-125-13	Sequence 13, Appli
37	869.8	99.3	873	9	US-10-830-825-13	Sequence 13, Appli
38	869.8	99.3	873	10	US-10-962-334-13	Sequence 13, Appli
39	869.8	99.3	873	10	US-10-962-235-13	Sequence 13, Appli
40	869.8	99.3	873	10	US-10-961-882-13	Sequence 13, Appli
41	869.8	99.3	873	10	US-10-961-828A-11	Sequence 11, Appli
42	869.8	99.3	873	10	US-10-303-161-11	Sequence 11, Appli
43	866.6	98.9	873	7	US-10-303-118-11	Sequence 11, Appli
44	866.6	98.9	873	7		
45	866.6	98.9	873	7		

# ALIGNMENTS

## RESULT 1

US-09-816-028A-8  
; Sequence 8, Application US/09816028A  
; Patent No. US20020042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19

US-09-816-028A-8

```

; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-816-028A-8

Query Match 100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 1.5e-137;
Matches: 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGAACCAAGTTTAAAGAAATGATTATTCAAGG 60
|||
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGAACCAAGTTTAAAGAAATGATTATTCAAGG 60
|||

QY 61 CTACCAAAATGATTTTCATGTATTAGATGTAATCAATTTTATTGTAAGATAAATCTAT 120

```

```
Db 61 CTACCAAAATGATTTTGATGTATTTAGATGTAAATCAATTTTATTTTGAAGATAAAATCTAT 120
Qy 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCAAAATCAAGATATGAGCCGAACTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCAAAATCAAGATATGAGCCGAACTAAATATGTGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTCTCT 300
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAA 360
Qy 361 TTTACGAAATTTTAAATCAAAAGAAATTTACCTCAGGGGCTATATGTGTGACGAGCC 420
Db 361 TTTACGAAATTTTAAATCAAAAGAAATTTACCTCAGGGGCTATATGTGTGACGAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAAAGAAATTTCTTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAAGAAATTTCTTTTAAACTAGCCCTGATTTTAAAT 540
Qy 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTTACAAATAAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660
Db 601 GAAAAAATTTTACAAATAAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATACAAAGAAATTTTAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATACAAAGAAATTTTAACTACACT 720
Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAATATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
```

## RESULT 2

```
US-10-303-161-8
; Sequence 8, Application US/10303161
; Publication No. US20030148459A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
```

```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3(alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:19
US-10-303-161-8
```

```
Query Match 100.0%; Score 876; DB 7; Length 876;
Best Local Similarity 100.0%; Pred. No. 1.5e-137; Indels 0; Gaps 0;
Matches 876; Conservative 0; Mismatches 0;

Qy 1 ATGAAAAAGTTATTATTCTGCGAAATGACCAAGTTTAAAGAAAAATTTGATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCTGCGAAATGACCAAGTTTAAAGAAAAATTTGATTATTCAAGG 60
Qy 61 CTACCAAAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCAAAATCAAGATATGAGCCGAACTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCAAAATCAAGATATGAGCCGAACTAAATATGTGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTCTCT 300
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAAATGCTTTATTTAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAAATGCTTTATTTAA 360
Qy 361 TTTACGAAATTTTAAATCAAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 420
Db 361 TTTACGAAATTTTAAATCAAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAAAGAAATTTCTTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAAGAAATTTCTTTTAAACTAGCCCTGATTTTAAAT 540
Qy 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTTACAAATAAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660
Db 601 GAAAAAATTTTACAAATAAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATACAAAGAAATTTTAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATACAAAGAAATTTTAACTACACT 720
Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAATATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATATAAATTTAAAGAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
```

QY 841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876  
|||||  
Db 841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876

## RESULT 3

US-10-303-118-8  
; Sequence 8, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Michel  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-118-8

Query Match 100.0%; Score 876; DB 7; Length 876;  
Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAAAATGATTTGATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120  
Db 61 CTACCAAAATGATTTGATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTTAATTTCTTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTTAATTTCTTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAATTAATGTTGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAATTAATGTTGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTGTAAACCTTTTACGATTTATTTGCT 300  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTGTAAACCTTTTACGATTTATTTGCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360  
QY 361 TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTTGTCAGTAGCC 420  
Db 361 TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTTGTCAGTAGCC 420  
QY 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATCGGTCA 480  
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATCGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
|||||  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGAGCTTTAGAAATTTCTA 600  
|||||  
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAACTTTACAAAAATAAACTATATTGCTTATGCTCTAATAGTCTTTTACAAATTTT 660  
|||||  
Db 601 GAAAAAACTTTACAAAAATAAACTATATTGCTTATGCTCTAATAGTCTTTTACAAATTTT 660  
QY 661 ATGAACTAGCGCAAAATTTAAATTTTAAATTTTATCATCAAGAAAAATAAATCACTACT 720  
|||||  
Db 661 ATGAACTAGCGCAAAATTTAAATTTTAAATTTTATCATCAAGAAAAATAAATCACTACT 720  
QY 721 AAGATATATCTCATACCTTCTAGTGAGCTTATCGAAATTTTCAAAAAATATTAATTTT 780  
|||||  
Db 721 AAGATATATCTCATACCTTCTAGTGAGCTTATCGAAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAATAAAAAATTAAGAAATGTTTATTACAAAGTTGATAAAGATCTTATTAAGATTA 840  
|||||  
Db 781 AAAAAATAAAAAATTAAGAAATGTTTATTACAAAGTTGATAAAGATCTTATTAAGATTA 840  
QY 841 CCTAGTGATATAAAGCATTATTTCAGAGGAAATAA 876  
|||||  
Db 841 CCTAGTGATATAAAGCATTATTTCAGAGGAAATAA 876

## RESULT 4

US-10-303-128-8  
; Sequence 8, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-128-8

Query Match 100.0%; Score 876; DB 7; Length 876;  
Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAAAATGATTTGATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120  
Db 61 CTACCAAAATGATTTGATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120

121 CTTGTAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180  
121 CTTGTAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180  
181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCGAATTAATATGTTCTTAAT 240  
181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCGAATTAATATGTTCTTAAT 240  
241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTCTT 300  
241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTCTT 300  
301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTTAA 360  
301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTTAA 360  
361 TTTACGAAATTTAATCAAAAGAAATTTACCTCAGGGTCTATATGTTGCGATGCC 420  
361 TTTACGAAATTTAATCAAAAGAAATTTACCTCAGGGTCTATATGTTGCGATGCC 420  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTTATCAAAATGGTCA 480  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTTATCAAAATGGTCA 480  
481 TCTTATGCTTTGATACCAAAATTTTAACTAGCCCTGATTTTAAAAAT 540  
481 TCTTATGCTTTGATACCAAAATTTTAACTAGCCCTGATTTTAAAAAT 540  
541 GATGCTCGACTATATCGGACATATGTAATAATACAGATATATAAGCTTTAGAA 600  
541 GATGCTCGACTATATCGGACATATGTAATAATACAGATATATAAGCTTTAGAA 600  
601 GAAAAAATTTAAATAAATAATATATGCTTATGCTTAAATGCTTTTACCAATTT 660  
601 GAAAAAATTTAAATAAATAATATATGCTTATGCTTAAATGCTTTTACCAATTT 660  
661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGCAAGAAAAATTAACCTAC 720  
661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGCAAGAAAAATTAACCTAC 720  
721 AAAGATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTT 780  
721 AAAGATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTT 780  
781 AAAAAATTAATAAATAAATAATTTTACAGTTGATAAAGATCTATTAGATTA 840  
781 AAAAAATTAATAAATAAATAATTTTACAGTTGATAAAGATCTATTAGATTA 840

RESULT 5  
US-10-303-134-8  
; Sequence 8, Application US/10303134  
; Publication No. US20030157657A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31

; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-134-8

Query Match 100.0%; Score 876; DB 7; Length 876;  
Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60  
QY 61 CTACCAAAATGATTTTATGATGATTTTATGATGATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTATGATGATTTTATGATGATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTTGTAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180  
Db 121 CTTGTAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCGAATTAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCGAATTAATATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTCTT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAAACTTTTACGATTTTCTT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
QY 361 TTTACGAAATTTAATCAAAAGAAATTTACCTCAGGGTCTATATGTTGCGATGCC 420  
Db 361 TTTACGAAATTTAATCAAAAGAAATTTACCTCAGGGTCTATATGTTGCGATGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTTATCAAAATGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTGATACCAAAATTTTAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTGATACCAAAATTTTAACTAGCCCTGATTTTAAAAAT 540  
QY 541 GATGCTCGACTATATCGGACATATGTAATAATACAGATATATAAGCTTTAGAA 600  
Db 541 GATGCTCGACTATATCGGACATATGTAATAATACAGATATATAAGCTTTAGAA 600  
QY 601 GAAAAAATTTAAATAAATAATTTTACAGTTGATAAAGATCTATTAGATTT 660  
Db 601 GAAAAAATTTAAATAAATAATTTTACAGTTGATAAAGATCTATTAGATTT 660  
QY 661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGCAAGAAAAATTAACCTAC 720  
Db 661 ATAGAACTAGGCGCAAAATTTAAATTTTAACTAGCAAGAAAAATTAACCTAC 720  
QY 721 AAAGATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTT 780  
Db 721 AAAGATATCTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTT 780  
QY 781 AAAAAATTAATAAATAAATAATTTTACAGTTGATAAAGATCTATTAGATTA 840  
Db 781 AAAAAATTAATAAATAAATAATTTTACAGTTGATAAAGATCTATTAGATTA 840





QY 121 CTTGGTAAATAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACACTAC 180  
Db 121 CTTGGTAAATAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACACTAC 180  
QY 181 ACTTTAAACACATTTAAATCCAAATCAAGAAATGAGCCGAACTAAATATGTTCTTAAT 240  
Db 181 ACTTTAAACACATTTAAATCCAAATCAAGAAATGAGCCGAACTAAATATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTAATTTCTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTAATTTCTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTTAATTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTTAATTTAAA 360  
QY 361 TTTACAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTACAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAAATTTGATTTTATCAAAATGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAATCAAGAAATTTTAAACCTTTTAAAGAAATTTAATTTAAA 540  
Db 481 TCTTATGCTTTTGATACCAATCAAGAAATTTTAAACCTTTTAAAGAAATTTAATTTAAA 540  
QY 541 GATCGCTCGCACTATATCGGCATAGTAAATATACAGATATAGTAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATATACAGATATAGTAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTTACAAATATAATTTATGCTTATGCTTCTTAATAGTCTTTTAGCAATTTT 660  
Db 601 GAAAAAATTTTACAAATATAATTTATGCTTATGCTTCTTAATAGTCTTTTAGCAATTTT 660  
QY 661 ATAGAACTAGCCCAATTTAAATTTTAACTTATCATCAAGAAAAATACTACACT 720  
Db 661 ATAGAACTAGCCCAATTTAAATTTTAACTTATCATCAAGAAAAATACTACACT 720  
QY 721 AAAGATATATCTCATACCTTTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATATTTAAT 780  
Db 721 AAAGATATATCTCATACCTTTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATATTTAAT 780  
QY 781 AAAAAATATAAATTAAGAAATTTTATTTAATCAAGTGTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTAAGAAATTTTATTTAATCAAGTGTGATAAAGATCTATTAAAGATTA 840

RESULT 8

US-10-820-536-8  
; Sequence 8, Application US/10820536  
; Publication No. US20040203103A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; PRIORITY FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406

; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3(alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CscII) from C. jejuni O:19  
US-10-820-536-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGGAACCAAGTTTAAAAAGAAATTTGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAATGGAACCAAGTTTAAAAAGAAATTTGATTATTTCAAGG 60  
QY 61 CTACCAAAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTACCAAAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
QY 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACACTAC 180  
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACACTAC 180  
QY 181 ACTTTAAAAATTTAAATCCAAATCAAGAAATATGAGCCGAACTAAATATGTTCTTAAT 240  
Db 181 ACTTTAAAAATTTAAATCCAAATCAAGAAATATGAGCCGAACTAAATATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTAATTTCTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTAATTTCTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAATTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTTAATTTAAA 360  
QY 361 TTTACAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTACAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAAATTTGATTTTATCAAAATGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAATCAAGAAATTTTAAACCTTTTAAAGAAATTTAATTTAAA 540  
Db 481 TCTTATGCTTTTGATACCAATCAAGAAATTTTAAACCTTTTAAAGAAATTTAATTTAAA 540  
QY 541 GATCGCTCGCACTATATCGGCATAGTAAATATACAGATATAGTAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATATACAGATATAGTAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTTACAAATATAATTTATGCTTATGCTTCTTAATAGTCTTTTAGCAATTTT 660  
Db 601 GAAAAAATTTTACAAATATAATTTATGCTTATGCTTCTTAATAGTCTTTTAGCAATTTT 660  
QY 661 ATAGAACTAGCCCAATTTAAATTTTAACTTATCATCAAGAAAAATACTACACT 720  
Db 661 ATAGAACTAGCCCAATTTAAATTTTAACTTATCATCAAGAAAAATACTACACT 720  
QY 721 AAAGATATATCTCATACCTTTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATATTTAAT 780  
Db 721 AAAGATATATCTCATACCTTTCTAGTAGGCTTTATGGAATAATTTTCAAAAAATATTTAAT 780  
QY 781 AAAAAATATAAATTAAGAAATTTTATTTAATCAAGTGTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTAAGAAATTTTATTTAATCAAGTGTGATAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAGCAATTATTTCAAGGAAATATA 876  
 Db 841 CCTAGTGATATAAGCAATTATTTCAAGGAAATATA 876

RESULT 9

US-10-845-408-8  
 ; Sequence 8, Application US/10845408  
 ; Publication No. US20040203112A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/845,408  
 ; CURRENT FILING DATE: 2004-05-12  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-845-408-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60  
 Db 1 ATGAAAAAGTTTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60  
 QY 61 CTACCAATGATTTTATGATGATTAATGATGATTAATGATGATTAATGATGATTAAT 120  
 Db 61 CTACCAATGATTTTATGATGATTAATGATGATTAATGATGATTAATGATGATTAAT 120  
 QY 121 CTGTGTAATAAATGCAAGCAGTGTGTTACACCCCTTAATTTCTTTGAGCAACTACTAC 180  
 Db 121 CTGTGTAATAAATGCAAGCAGTGTGTTACACCCCTTAATTTCTTTGAGCAACTACTAC 180  
 QY 181 ACTTTAAACCAATTTAATCCAAATCAAGATATGACCGCAACTAATTTATGTTCTTAAT 240  
 Db 181 ACTTTAAACCAATTTAATCCAAATCAAGATATGACCGCAACTAATTTATGTTCTTAAT 240  
 QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300  
 Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300  
 QY 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGATTTTAAAGATTTTAAAG 360  
 Db 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGATTTTAAAGATTTTAAAG 360  
 QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTCAGTACGAGG 420  
 Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTCAGTACGAGG 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATCGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATCGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 QY 541 GATCGCTCGCACTATATCGGCATAGTAAATAATCAGATATATAAAGCTTTAGAAATTTCTA 600  
 Db 541 GATCGCTCGCACTATATCGGCATAGTAAATAATCAGATATATAAAGCTTTAGAAATTTCTA 600  
 QY 601 GAAAAAACTTTACAAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACCAAAATTTT 660  
 Db 601 GAAAAAACTTTACAAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACCAAAATTTT 660  
 QY 661 ATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAGAAAAATAAATTAACACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAGAAAAATAAATTAACACT 720  
 QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
 Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
 QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAGATTATA 840  
 Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAGATTATA 840  
 QY 841 CCTAGTGATATAAGCAATTATTTCAAGGAAATATA 876  
 Db 841 CCTAGTGATATAAGCAATTATTTCAAGGAAATATA 876

RESULT 10

US-10-845-412-8  
 ; Sequence 8, Application US/10845412  
 ; Publication No. US20040203113A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/845,412  
 ; CURRENT FILING DATE: 2004-05-12  
 ; PRIOR APPLICATION NUMBER: US/10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-845-412-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60  
 Db 1 ATGAAAAAGTTTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60  
 QY 61 CTACCAATGATTTTATGATGATTAATGATGATTAATGATGATTAATGATGATTAAT 120

Db 61 CTACCAATGATTTGATGTAATTTAGATGTAATCAATTTATTTTGAAGATAAATACAT 120  
Qy 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATAC 180  
Db 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATAC 180  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGACCGCACTAATTTATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGACCGCACTAATTTATGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACAGATTATTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACAGATTATTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
Qy 361 TTTCAACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTCAACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Qy 541 GATCGCTCGCATATATCGGACATAGTAAATATACAGATATAAAGCTTTTGAATTTCTA 600  
Db 541 GATCGCTCGCATATATCGGACATAGTAAATATACAGATATAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATAAAATCTATATGCTTATGCTTATAGTCTTTTACGAAATTT 660  
Db 601 GAAAAAATTTACAAAATAAAATCTATGCTTATGCTTATAGTCTTTTACGAAATTT 660  
Qy 661 ATAGAATCTAGCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAAATACACACT 720  
Db 661 ATAGAATCTAGCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAAATACACACT 720  
Qy 721 AAAGATATACATACACTCTTAGTGGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780  
Db 721 AAAGATATACATACACTCTTAGTGGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780  
Qy 781 AAAAAATATAAATTAAGAAATGTTTATTAAGATTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTAAGAAATGTTTATTAAGATTTGATAAAGATCTATTAAAGATTA 840

RESULT 11

US-10-846-219-8  
; Sequence 8, Application US/10846219  
; Publication No. US20040219638A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/846,219  
; CURRENT FILING DATE: 2004-05-14  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-846-219-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
Best Local Similarity 100.0%; Pred. No. 1.5e-137; Mismatches 0; Gaps 0;  
Matches 876; Conservative 0; Indels 0;  
Qy 1 ATGAAAAAGTTATTTATCTCGAAATGGACCAAGTTTAAAGAAAATTTGATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTTATCTCGAAATGGACCAAGTTTAAAGAAAATTTGATTTCAAGG 60  
Qy 61 CTACCAAAATGATTTTGATGTATTTAGATGTATCAATTTTATTTGGAAGATAAATACAT 120  
Db 61 CTACCAAAATGATTTTGATGTATTTAGATGTATCAATTTTATTTGGAAGATAAATACAT 120  
Qy 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATAC 180  
Db 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATAC 180  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGACCGCACTAATTTATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGACCGCACTAATTTATGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
Qy 361 TTTCAACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTCAACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Qy 541 GATCGCTCGCATATATCGGACATAGTAAATATACAGATATAAAGCTTTTGAATTTCTA 600  
Db 541 GATCGCTCGCATATATCGGACATAGTAAATATACAGATATAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAAATAAAATCTATATGCTTATGCTTATAGTCTTTTACGAAATTT 660  
Db 601 GAAAAAATTTACAAAATAAAATCTATATGCTTATGCTTATAGTCTTTTACGAAATTT 660  
Qy 661 ATAGAATCTAGCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAAATACACACT 720  
Db 661 ATAGAATCTAGCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAAATACACACT 720  
Qy 721 AAAGATATACATACACTCTTAGTGGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780  
Db 721 AAAGATATACATACACTCTTAGTGGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780  
Qy 781 AAAAAATATAAATTAAGAAATGTTTATTAAGATTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTAAGAAATGTTTATTAAGATTTGATAAAGATCTATTAAAGATTA 840

Db 781 AAAAAAATAAATAAAGAAATGTTATTACAGTTCATAAAGATCTATTAAAGTTA 840  
 QY 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876  
 Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876

RESULT 12

US-10-821-604-8  
 ; Sequence 8, Application US/10821604  
 ; Publication No. US20040229263A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/821,604  
 ; CURRENT FILING DATE: 2004-04-08  
 ; PRIOR APPLICATION NUMBER: 10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-821-604-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATCTCGAAATGCAACCAAGTTTAAAGAAATTCGATTATTCACAGG 60  
 Db 1 ATGAAAAAGTTATTATCTCGAAATGCAACCAAGTTTAAAGAAATTCGATTATTCACAGG 60  
 QY 61 CTACCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 Db 61 CTACCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 QY 121 CTGTGTAATAATGCAAGCAAGTGTTCACCCCTAAATTCCTCTTTGAGCAAACTACTAC 180  
 Db 121 CTGTGTAATAATGCAAGCAAGTGTTCACCCCTAAATTCCTCTTTGAGCAAACTACTAC 180  
 QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGCACTAATATGTTCTTAAT 240  
 Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGCACTAATATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300  
 Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300  
 QY 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAAGAAATTTAAAG 360  
 Db 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAAGAAATTTAAAG 360  
 QY 361 TTTACGAAATTTTATTCATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
 Db 361 TTTACGAAATTTTATTCATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATCGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATCGGTCA 480  
 QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAT 540  
 Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGAT 540  
 QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA 600  
 Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA 600  
 QY 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTCTCTATAGTCTTTTAAAGAAATTTT 660  
 Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTCTCTATAGTCTTTTAAAGAAATTTT 660  
 QY 661 ATAGAACTAGCGCAATTTTAAATTTCAAAATTTTATCATACAGAAATAAATTAACACTAC 720  
 Db 661 ATAGAACTAGCGCAATTTTAAATTTCAAAATTTTATCATACAGAAATAAATTAACACTAC 720  
 QY 721 AAAGATATATCTCATACCTCTCTAGTGAGCTTTATCGAAAATTTTCAAAAAATATTAATTTT 780  
 Db 721 AAAGATATATCTCATACCTCTCTAGTGAGCTTTATCGAAAATTTTCAAAAAATATTAATTTT 780  
 QY 781 AAAAAAATAAATAAAGAAATGTTTATTAACAAGTTGATATAAGATCTATTAAGATTA 840  
 Db 781 AAAAAAATAAATAAAGAAATGTTTATTAACAAGTTGATATAAGATCTATTAAGATTA 840  
 QY 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876  
 Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAATAA 876

RESULT 13

US-10-847-983-8  
 ; Sequence 8, Application US/10847983  
 ; Publication No. US20040229272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/847,983  
 ; CURRENT FILING DATE: 2004-05-17  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-847-983-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATCTCGAAATGCAACCAAGTTTAAAGAAATTTGATTTTCAAGG 60  
 Db 1 ATGAAAAAGTTATTATCTCGAAATGCAACCAAGTTTAAAGAAATTTGATTTTCAAGG 60

```
Qy 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAATGCAAGCAGTGTCTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAATGCAAGCAGTGTCTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAATTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAATTAATTTATGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATATTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATATTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360
Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTCAAAAATGGGTCA 480
Qy 481 TCCTATGCTTTTATACCAACAAAGAAATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 540
Db 481 TCCTATGCTTTTATACCAACAAAGAAATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 540
Qy 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTTCAAAAATAAATCTATTTGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTTCAAAAATAAATCTATTTGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACTAC 720
Db 661 ATAGAACTAGGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACTAC 720
Qy 721 AAAGATATCTATACCTTTAGTGGCTTATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 780
Db 721 AAAGATATCTATACCTTTAGTGGCTTATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 780
```

## RESULT 14

```
US-10-821-573-8
; Sequence 8, Application US/10821573
; Publication No. US20040229313A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; PRIORITY FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
```

```
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-821-573-8
```

```
Query Match 100.0%; Score 876; DB 9; Length 876;
Best Local Similarity 100.0%; Pred. No. 1.5e-137;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ATGAAAAAGTTTATTATTCCTCGAAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTTATTATTCCTCGAAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Qy 61 CTACCAATGATTTTGTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAAATCAAGAAATGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATGACCGAACTAATTTATGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATATTTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360
Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCCTATGCTTTTATACCAACAAAGAAATTTTAAACAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCCTATGCTTTTATACCAACAAAGAAATTTTAAACAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTTCAAAAATAAATCTATTTGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTTCAAAAATAAATCTATTTGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACTAC 720
Db 661 ATAGAACTAGGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATTAACACTAC 720
Qy 721 AAAGATATCTATACCTTTAGTGGCTTATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 780
Db 721 AAAGATATCTATACCTTTAGTGGCTTATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 780
```

QY 781 AAAAAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 781 AAAAAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
 QY 841 CCTAGTGATATAAAGCAATATTTCAGAGGAAAAATAA 876  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 841 CCTAGTGATATAAAGCAATATTTCAGAGGAAAAATAA 876

RESULT 15  
 US-10-850-807-8  
 ; Sequence 8, Application US/10850807  
 ; Publication No. US20040259140A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/850,807  
 ; CURRENT FILING DATE: 2004-05-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-850-807-8

Query Match 100.0%; Score 876; DB 9; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-137;  
 Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTGCTGGAATGCAACCAAGTTTAAAGAAATTCGATTATTCACAGG 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 1 ATGAAAAAGTTATTATTGCTGGAATGCAACCAAGTTTAAAGAAATTCGATTATTCACAGG 60  
 QY 61 CTACCAATGATTTTGTATGTTATAGATGTAATCAATTTTATTTCAGATATAATCTAT 120  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 61 CTACCAATGATTTTGTATGTTATAGATGTAATCAATTTTATTTCAGATATAATCTAT 120  
 QY 121 CTGTGTAATAAATGCAAGCAGTGTTCACCCCTAAATTCCTCTTCAGCAATACTAC 180  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 121 CTGTGTAATAAATGCAAGCAGTGTTCACCCCTAAATTCCTCTTCAGCAATACTAC 180  
 QY 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGCCGAATTAATGTTCTTAAT 240  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGCCGAATTAATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTATTTCCT 300  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTATTTCCT 300  
 QY 301 GATGCTCAATTTGGGATGATTTTTTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTAA 360  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 301 GATGCTCAATTTGGGATGATTTTTTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTAA 360  
 QY 361 TTTCACGAATTTATTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCTAGTAC 420  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 361 TTTCACGAATTTATTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCTAGTAC 420

QY 421 ATAGCCCTAGGATCAAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 421 ATAGCCCTAGGATCAAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
 QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540  
 QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
 QY 601 GAAAAAACTTACAAAAATAAAACTATATTGCTTATGCTCCTTAATAGTCTTTTACGCAATTTT 660  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 601 GAAAAAACTTACAAAAATAAAACTATATTGCTTATGCTCCTTAATAGTCTTTTACGCAATTTT 660  
 QY 661 ATAGAACTAGCGCCCAAAATTTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 661 ATAGAACTAGCGCCCAAAATTTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720  
 QY 721 AAAGATATACTCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATATTATTAATTTT 780  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 721 AAAGATATACTCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATATTATTAATTTT 780  
 QY 781 AAAAAAATAAATAAAGAAATGTTTATTACAAGTTTGATAAAGATCTATTAAAGATTA 840  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 781 AAAAAAATAAATAAAGAAATGTTTATTACAAGTTTGATAAAGATCTATTAAAGATTA 840  
 QY 841 CCTAGTGATATAAAGCAATTTTTCAGAGGAAAAATAA 876  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 841 CCTAGTGATATAAAGCAATTTTTCAGAGGAAAAATAA 876

Search completed: July 19, 2006, 15:59:48  
 Job time : 1000.62 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: July 19, 2006, 13:47:11 ; Search time 27.928 Seconds  
(without alignments)  
912.038 Million cell updates/sec

Title: US-10-734-719-9  
Perfect score: 1563  
Sequence: 1 MKKVIIAGNGSLKEIDYSR.....KLIKDLLRLPSDIKHVFKGK 291

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Issued Patents AA.\*
- 1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5\_COMB.pep.\*
  - 2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6\_COMB.pep.\*
  - 3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7\_COMB.pep.\*
  - 4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H\_COMB.pep.\*
  - 5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/PCTUS\_COMB.pep.\*
  - 6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE\_COMB.pep.\*
  - 7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1563	100.0	291	2	US-09-495-406-9
2	1563	100.0	291	2	US-09-816-028A-9
3	1563	100.0	291	2	US-10-303-162-9
4	1563	100.0	291	2	US-10-303-134-9
5	1563	100.0	291	2	US-10-303-118-9
6	1563	100.0	291	2	US-10-303-128-9
7	1563	100.0	291	3	US-10-735-419-9
8	1559	99.7	291	2	US-09-816-028A-14
9	1559	99.7	291	2	US-10-303-162-14
10	1559	99.7	291	2	US-10-303-134-14
11	1559	99.7	291	2	US-10-303-118-14
12	1559	99.7	291	2	US-10-303-128-14
13	1559	99.7	291	3	US-10-735-419-14
14	1557	99.6	291	2	US-09-816-028A-12
15	1557	99.6	291	2	US-10-303-162-12
16	1557	99.6	291	2	US-10-303-134-12
17	1557	99.6	291	2	US-10-303-118-12
18	1557	99.6	291	2	US-10-303-128-12
19	1557	99.6	291	3	US-10-735-419-12
20	1528	97.8	291	2	US-09-495-406-5
21	1528	97.8	291	2	US-09-816-028A-5
22	1528	97.8	291	2	US-10-303-162-5
23	1528	97.8	291	2	US-10-303-134-5
24	1528	97.8	291	2	US-10-303-118-5
25	1528	97.8	291	2	US-10-303-128-5
26	1528	97.8	291	3	US-10-735-419-5

27	1525	97.6	291	2	US-09-495-406-3	Sequence 3, Appli
28	1525	97.6	291	2	US-09-816-028A-3	Sequence 3, Appli
29	1525	97.6	291	2	US-10-303-162-3	Sequence 3, Appli
30	1525	97.6	291	2	US-10-303-134-3	Sequence 3, Appli
31	1525	97.6	291	2	US-10-303-118-3	Sequence 3, Appli
32	1525	97.6	291	2	US-10-303-128-3	Sequence 3, Appli
33	1525	97.6	291	3	US-10-735-419-3	Sequence 3, Appli
34	1514	96.9	291	2	US-09-495-406-7	Sequence 7, Appli
35	1514	96.9	291	2	US-09-816-028A-7	Sequence 7, Appli
36	1514	96.9	291	2	US-10-303-162-7	Sequence 7, Appli
37	1514	96.9	291	2	US-10-303-134-7	Sequence 7, Appli
38	1514	96.9	291	2	US-10-303-118-7	Sequence 7, Appli
39	1514	96.9	291	2	US-10-303-128-7	Sequence 7, Appli
40	1514	96.9	291	3	US-10-735-419-7	Sequence 7, Appli
41	804	51.4	294	2	US-09-495-406-10	Sequence 10, Appl
42	804	51.4	294	2	US-09-816-028A-10	Sequence 10, Appl
43	804	51.4	294	2	US-10-303-162-10	Sequence 10, Appl
44	804	51.4	294	2	US-10-303-134-10	Sequence 10, Appl
45	804	51.4	294	2	US-10-303-118-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-495-406-9  
; Sequence 9, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; US-09-495-406-9

Query Match 100.0%; Score 1563; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 1.9e-155;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGSLKEIDYSRLPNDPFDVRCNQYFEDKYILGKKCAVFTPNPFQYY 60  
DB 1 MKKVIIAGNGSLKEIDYSRLPNDPFDVRCNQYFEDKYILGKKCAVFTPNPFQYY 60  
QY 61 TLKHLIQNEYETELIMCSNNOAHLENENFVKTDFYDPAHLGYDFFKOLKEFNAYFK 120  
DB 61 TLKHLIQNEYETELIMCSNNOAHLENENFVKTDFYDPAHLGYDFFKOLKEFNAYFK 120  
QY 121 FHEIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 FHEIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYIGHKNDIKALEFLEKTYKIKLCLCPNSLLANFIELAPNLSNFIIEKKNYT 240  
DB 181 DRSHYIGHKNDIKALEFLEKTYKIKLCLCPNSLLANFIELAPNLSNFIIEKKNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291

RESULT 2

```
US-09-816-028A-9
; Sequence 9, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-816-028A-9

Query Match      100.0%; Score 1563; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.9e-155;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRLPNDVDVFRNCQFYFEDKYILGKKCKAVFTYTPNFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVDVFRNCQFYFEDKYILGKKCKAVFTYTPNFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180
Db 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 3
US-10-303-162-9
; Sequence 9, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-162-9

Query Match      100.0%; Score 1563; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.9e-155;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRLPNDVDVFRNCQFYFEDKYILGKKCKAVFTYTPNFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVDVFRNCQFYFEDKYILGKKCKAVFTYTPNFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180
Db 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 4
US-10-303-134-9
; Sequence 9, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-134-9

Query Match      100.0%; Score 1563; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.9e-155;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRLPNDVDVFRNCQFYFEDKYILGKKCKAVFTYTPNFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVDVFRNCQFYFEDKYILGKKCKAVFTYTPNFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180
Db 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
```



Db 61 TLKHLIQNEYETELIMCSNYNQAHLNENFVKTFYDFPDALHGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

## RESULT 5

US-10-303-118-9  
; Sequence 9, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-118-9

Query Match 100.0%; Score 1563; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 1.9e-155;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLGKCKKAVFTPNPFEEQYY 60  
Db 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLGKCKKAVFTPNPFEEQYY 60  
QY 61 TLKHLIQNEYETELIMCSNYNQAHLNENFVKTFYDFPDALHGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQNEYETELIMCSNYNQAHLNENFVKTFYDFPDALHGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

## RESULT 6

US-10-303-128-9

; Sequence 9, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-128-9

Query Match 100.0%; Score 1563; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 1.9e-155;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLGKCKKAVFTPNPFEEQYY 60  
Db 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLGKCKKAVFTPNPFEEQYY 60  
QY 61 TLKHLIQNEYETELIMCSNYNQAHLNENFVKTFYDFPDALHGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQNEYETELIMCSNYNQAHLNENFVKTFYDFPDALHGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

## RESULT 7

US-10-735-419-9  
; Sequence 9, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRF
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-735-419-9

Query Match          100.0%; Score 1563; DB 3; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.9e-155;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLGKCKKAVFYTPNFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLGKCKKAVFYTPNFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120

QY 121 FHEIYFNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
DB 121 FHEIYFNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
US-09-816-028A-14
; Sequence 14, Application US/09816028A
; Patent No. 669705
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRF
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-09-816-028A-14

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 4.9e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLGKCKKAVFYTPNFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLGKCKKAVFYTPNFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120

QY 121 FHEIYFNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
DB 121 FHEIYFNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 10
US-10-734-134-14
; Sequence 14, Application US/10303134
```

```
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
; US-10-303-134-14

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 4.9e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRPNDFVFRCNQFYFEDKYLGGKCAVFTYTPNFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVFRCNQFYFEDKYLGGKCAVFTYTPNFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFDYFPPDAHLGYDFFKQLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFDYFPPDAHLGYDFFKQLKEFNAYFK 120

Qy 121 FHEIYNQRTSGVMCAVAIALGYKEIYLSGIDFYONGSSVAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYNQRTSGVMCAVAIALGYKEIYLSGIDFYONGSSVAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 11
US-10-303-118-14
; Sequence 14, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
; US-10-303-118-14

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 4.9e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRPNDFVFRCNQFYFEDKYLGGKCAVFTYTPNFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVFRCNQFYFEDKYLGGKCAVFTYTPNFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFDYFPPDAHLGYDFFKQLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFDYFPPDAHLGYDFFKQLKEFNAYFK 120

Qy 121 FHEIYNQRTSGVMCAVAIALGYKEIYLSGIDFYONGSSVAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYNQRTSGVMCAVAIALGYKEIYLSGIDFYONGSSVAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 12
US-10-303-128-14
; Sequence 14, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
; US-10-303-128-14

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 4.9e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRPNDFVFRCNQFYFEDKYLGGKCAVFTYTPNFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVFRCNQFYFEDKYLGGKCAVFTYTPNFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFDYFPPDAHLGYDFFKQLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFDYFPPDAHLGYDFFKQLKEFNAYFK 120

Qy 121 FHEIYNQRTSGVMCAVAIALGYKEIYLSGIDFYONGSSVAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYNQRTSGVMCAVAIALGYKEIYLSGIDFYONGSSVAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
```

```

Db 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFDYFPDHAHLGYDFKQLKEFNAYFK 120
QY 121 FHEIYNQRTSGVVMCAVAIALGKEYIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFKN 180
Db 121 FHEIYNQRTSGVVMCAVAIALGKEYIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFKN 180
QY 181 DRSHYGHKNTDIAKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYGHKNTDIAKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKGK 291

```

```

RESULT 13
US-10-735-419-14
; Sequence 14, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (Cstii) from C. jejuni O:36
US-10-735-419-14

```

```

Query Match 99.7%; Score 1559; DB 3; Length 291;
Best Local Similarity 99.7%; Pred. No. 4.9e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRQNFYEDKYYLGKCKKAVFYTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRQNFYEDKYYLGKCKKAVFYTPNFFEQYY 60
QY 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFDYFPDHAHLGYDFKQLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFDYFPDHAHLGYDFKQLKEFNAYFK 120
QY 121 FHEIYNQRTSGVVMCAVAIALGKEYIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFKN 180
Db 121 FHEIYNQRTSGVVMCAVAIALGKEYIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFKN 180
QY 181 DRSHYGHKNTDIAKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYGHKNTDIAKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKGK 291

```

```

RESULT 14
US-09-816-028A-12

```

```

; Sequence 12, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (Cstii) from C. jejuni O:4
US-09-816-028A-12

Query Match 99.6%; Score 1557; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 8e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRQNFYEDKYYLGKCKKAVFYTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRQNFYEDKYYLGKCKKAVFYTPNFFEQYY 60
QY 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFDYFPDHAHLGYDFKQLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFDYFPDHAHLGYDFKQLKEFNAYFK 120
QY 121 FHEIYNQRTSGVVMCAVAIALGKEYIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFKN 180
Db 121 FHEIYNQRTSGVVMCAVAIALGKEYIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFKN 180
QY 181 DRSHYGHKNTDIAKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYGHKNTDIAKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKGK 291

```

```

RESULT 15
US-10-303-162-12
; Sequence 12, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12

```

```
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-162-12

Query Match      99.8%; Score 1557; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 8e-155;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MKKVIIAGNGPSLKEIDYSLPNDFDVFRCNQFYFEDKYVLGKKCKAVPYTPNFFFEQYY 60
Db      1 MKKVIIAGNGPSLKEIDYSLPNDFDVFRCNQFYFEDKYVLGKKCKAVPYTPGFFFEQYY 60

QY      61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYFPDAHLGYDFFKQKKEFNAYFK 120
Db      61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYFPDAHLGYDFFKQKKEFNAYFK 120

QY      121 FHEIYFNQRIITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 180
Db      121 FHEIYFNQRIITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 180

QY      181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPSNLLANFIELAPNLNSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPSNLLANFIELAPNLNSNFIIOEKNNYT 240

QY      241 KDILIPSSSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db      241 KDILIPSSSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
```

Search completed: July 19, 2006, 13:50:05  
Job time : 28.928 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 14:13:57 ; Search time 13.2159 Seconds  
(without alignments)  
1269.788 Million cell updated/sec

Title: US-10-734-719-9  
Perfect score: 1563  
Sequence: 1 MKVLIAGNPGSLKIDYSR.....KLIKLLRLPSDIKHYFKGK 291

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 208217 seqs, 57668156 residues

Total number of hits satisfying chosen parameters: 208217

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA\_New.\*  
1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97.5	6.2	496	6	US-10-471-571A-3348
2	97.5	6.2	651	6	US-10-471-571A-252
3	96	6.1	701	6	US-10-471-571A-1902
4	95.5	6.1	468	6	US-10-471-571A-5412
5	92.5	5.9	955	6	US-10-471-571A-708
6	91.5	5.9	437	7	US-11-293-697-4502
7	88	5.6	512	7	US-11-056-355B-91714
8	88	5.6	512	7	US-11-056-355B-95470
9	88	5.6	530	7	US-11-056-355B-91713
10	88	5.6	530	7	US-11-056-355B-95469
11	88	5.6	659	7	US-11-056-355B-91712
12	88	5.6	659	7	US-11-056-355B-95468
13	87	5.6	425	6	US-10-953-349-10851
14	87	5.6	429	6	US-10-953-349-10850
15	87	5.6	1337	7	US-11-257-500-29
16	86.5	5.5	863	6	US-10-527-411-36
17	86.5	5.5	1129	6	US-10-527-411-54
18	86.5	5.5	1129	6	US-10-527-411-52
19	86	5.5	396	6	US-10-449-902-50660
20	86	5.5	971	6	US-10-528-563-7
21	85.5	5.5	965	7	US-11-178-560-3
22	85.5	5.5	1143	6	US-10-630-629-1
23	85.5	5.5	1516	7	US-11-257-500-23
24	85	5.4	528	6	US-10-449-902-44124
25	84.5	5.4	965	7	US-11-253-453-3

26	84	5.4	452	7	US-11-199-489A-124	Sequence 124, App
27	84	5.4	488	7	US-11-056-355B-48594	Sequence 48594, A
28	84	5.4	533	7	US-11-056-355B-48593	Sequence 48593, A
29	84	5.4	601	7	US-11-056-355B-48592	Sequence 48592, A
30	83.5	5.3	476	6	US-10-449-902-32670	Sequence 32670, A
31	83.5	5.3	506	6	US-10-449-902-52094	Sequence 52094, A
32	83	5.3	301	6	US-10-471-571A-1484	Sequence 1484, Ap
33	83	5.3	877	6	US-10-527-411-157	Sequence 157, App
34	83	5.3	887	6	US-10-471-571A-722	Sequence 722, App
35	82.5	5.3	188	6	US-10-527-411-161	Sequence 161, App
36	82.5	5.3	863	6	US-10-527-411-38	Sequence 38, Appl
37	82.5	5.3	972	7	US-11-253-453-1	Sequence 1, Appli
38	82.5	5.3	972	7	US-11-253-089-52	Sequence 52, Appl
39	82.5	5.3	1127	6	US-10-527-411-50	Sequence 50, Appl
40	82.5	5.3	1129	6	US-10-527-411-48	Sequence 48, Appl
41	82	5.2	348	6	US-10-953-349-10852	Sequence 10852, A
42	82	5.2	456	7	US-11-056-355B-79879	Sequence 79879, A
43	82	5.2	478	7	US-11-056-355B-79878	Sequence 79878, A
44	82	5.2	521	7	US-11-056-355B-79877	Sequence 79877, A
45	82	5.2	864	7	US-11-251-208-77	Sequence 77, Appl

## ALIGNMENTS

### RESULT 1

US-10-471-571A-3348  
; Sequence 3348, Application US/10471571A  
; Publication No. US20060115490A1

; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA

; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927WO

; CURRENT APPLICATION NUMBER: US/10/471,571A  
; CURRENT FILING DATE: 2003-09-12

; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03

; SEQ ID NO 3348  
; LENGTH: 496

; TYPE: PRT  
; ORGANISM: Staphylococcus aureus

; FEATURE:  
; NAME/KEY: MISC FEATURE

; LOCATION: (1)..(496)  
; OTHER INFORMATION: hypothetical protein

US-10-471-571A-3348

Query Match 6.2%; Score 97.5; DB 6; Length 496;  
Best Local Similarity 20.0%; Pred. No. 1.1;  
Matches 65; Conservative 51; Mismatches 114; Indels 95; Gaps 16;

QY	12	SLKEIDYSLPNDPFCNQF----	YFEDKY-----	LGKKCAVFTPNFFF	56
DB	90	TLKHVENS---HDFRIYDQERFLWYAHFQDPKRYLDYVNHFDQRKVKRDFYDVRGFL	146		
QY	57	-----EQYILKHLIQEYETELIMCSNYQAHLNENFVKTF-----	YDFPDHGLG	105	
DB	147	SGSRILVDKQOTLCEFFYNPEGDTKLEKPYSGDKGPEVQKIIIVYANKQYFFNNETELG	206		
QY	106	YDFKOLKEFNAYFKFHEIVNQRTISGVVYMC-----	AVATLGYKEILSGID---	154	
DB	207	AFPIKOL-----YQMGDLFFSDR---NVYTAIFNUTPESIPVAVLHSTHINKNIDALD	257		
QY	155	-----FYQNGSSY---AFDTKQENLLKLAPDFKNDRSHYI-----	GHSKNTDI	194	
DB	258	SSPEKNVYKAMPENLSRYRAIIIVSTEQKL-----	DVEKRINHPTIPVNVIPVGYSITDT	312	
QY	195	KALEFLEKTYK-IKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYTKDILIPSS	-----	249	
DB	313	PVQTLTDQRSVKLISVARYSPEKHOQIIEIKRLVS-----	YVPKIELHMY	358	

QY 250 AYGFESKNIN--FKKIKIKENVYK 272  
Db 359 GFGSESKNLNELIQYGLNHHVLR 383

RESULT 2  
US-10-471-571A-252  
; Sequence 252, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927WO  
; CURRENT APPLICATION NUMBER: US/10/471, 571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 252  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)..(651)  
; OTHER INFORMATION: hypothetical ABC  
US-10-471-571A-252

Query Match 6.2%; Score 97.5; DB 6; Length 651;  
Best Local Similarity 22.9%; Pred. No. 1.5;  
Matches 51; Conservative 29; Mismatches 66; Indels 77; Gaps 11;

QY 14 KEIDY-SRLPNDF--DVFR-----CNQYFEDKYVLGKKCAVFTNPFPEQY-- 59  
Db 112 EETDWSKHANDYSDTYKTMRSYSLNFEQLEGQYRESKIYTVLYGLNFSSEDFNK 171  
QY 60 -----YTLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDPPDAHLGYD 107  
Db 172 PINDFSGQKTRLSLAQMLLN--EPDLULLDE-----PTNHLDL 209  
QY 108 FFKOLKEFNAYK-----FHEIYNQRTSGVYMCVAIAIALGYKEIYLSGI-DYQNGS 160  
Db 210 TTKWLDELYRFKGAIVAIISHDRYFLDKIVTIQY-----DVALGDVYKRVGNVEEPIQORD 265  
QY 161 SYAFDTKQENLLKLPDFKNDRSHVIGHSKNTDIALKLE-ELEK 202  
Db 266 LYYQKRWQY-----ESQAEIKRLFTFVEK 291

RESULT 3  
US-10-471-571A-1902  
; Sequence 1902, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927WO  
; CURRENT APPLICATION NUMBER: US/10/471, 571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 1902  
; LENGTH: 701  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)..(701)  
; OTHER INFORMATION: hypothetical protein  
US-10-471-571A-1902

Query Match 6.1%; Score 96; DB 6; Length 701;  
Best Local Similarity 19.7%; Pred. No. 2.3;  
Matches 71; Conservative 46; Mismatches 92; Indels 152; Gaps 19;

QY 15 EIDYSRLP-----NDPDPFRNQFYFEDKYVLGKKCAVFTNPF--EQ 58  
Db 324 DINFEHLPRVVFIDDIHDIETISQTN-----YNLLNRCFEKLFERNIGLAIKSTQQ 376  
QY 59 YVTLKHLI-----QNOEYETE-----LIMCSYNOA-----HLNEN-----F 91  
Db 377 FETMKQLIITFLQGNQDYKTSKKLVKFMVFCNSMTABEIHLCIKKNKEIKYSVT 436  
QY 92 VKTFYDYPDAHLGYDFFKOLKEFNAYF-----KHF--EIFYNQRTS 132  
Db 437 VDGFLFETYSTVEQVYDMQRLK-FHYVFIDIENSKTATHLITKNQHVQTDTHFEQ----- 491  
QY 133 GVTMCVAIAIALGYKEIYL-SGIDFYQ-----NGSSYAFDTKQENLLK----- 173  
Db 492 -----YKFLDSGISSTQFVYNNLSVSGFKYTDGKNPIQLSDIVYHIAL 539  
QY 174 ----LAPDFKNDRSHVIG-HSKNTDIALKLEFEKTYKIKLYCLCPNSLLANFIELAPNL 227  
Db 540 RYGGISYQLDDHSNYISLYNKYGSPLPLMHLKYQFR-----PFVNDEIEI 586  
QY 228 NSNFIIQENNYTKOILIPSEAYGKFSKNINFKKIKIKENVYVYKLIKDLRLPSDIKHY 287  
Db 587 TNNYVLSRKDN-----NYHF-----LLFNKIND--RYMSDVVKQD 618  
QY 288 F 288  
Db 619 F 619

RESULT 4  
US-10-471-571A-5412  
; Sequence 5412, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927WO  
; CURRENT APPLICATION NUMBER: US/10/471, 571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 5412  
; LENGTH: 468  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-5412

Query Match 6.1%; Score 95.5; DB 6; Length 468;  
Best Local Similarity 21.1%; Pred. No. 1.5;  
Matches 83; Conservative 49; Mismatches 113; Indels 149; Gaps 21;

QY 3 KVIAGNGSLKEIDYSRLPNDPDPFRNQFYFEDKYVL--GKKCAVFTNPFPEQY 60  
Db 109 EYVLSG---ELEKIVKHLFENKFDV-----EYGFYILFIKPIFEIVFSSNYVKEFN 157  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLENE-NFVKTFYDYPDAHLGYDFEKKOLK----- 113  
Db 158 T---KTESSEYRESLDSFNEILMKYLENEINFKEVWKYID-----ELFKLVNQVKLLI 208  
QY 114 -----EFNAYFKFHEIYFNQRTSGVYMCVAIAIALGYKEIY----- 149  
Db 209 HKKIESLISSYDKISDDCSMYSKVFEQYKKNVLDKVY-----YSEVYKSDNLI 259  
QY 150 -----LSGIDFYQNGSSYAFDTKQENLLKLPADFKNDRSHYIGHS-----KNTDIAL 197  
Db 260 KVFSEDVWNIIMGYKKYNEY-FTRDDTTTKIT-----NLYICQSFTEGTRNVIDRIG 311

QY 198 EFLKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYTKDIL-----IPS 247  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 248 SE-----AYGKFSKNIF-----KKIK--IKEN-----VYKLI 274  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 367 KEYIDLSEISLITFFKPEYKQHVNEFNLSLSSKIKNLAKENDGIIYINFKDQVNYQLI 426  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 275 -----KDLRLPS-DIKHYPK 289  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 427 SEGNTFFSAKPYLHCVFNKDILNMDKVDIENFFK 460  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

RESULT 5  
US-10-471-571A-708  
; Sequence 708, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927W0  
; CURRENT APPLICATION NUMBER: US/10/471,571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: Seqwin99, version 1.03  
; SEQ ID NO 708  
; LENGTH: 955  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-708

Query Match 5.9%; Score 92.5; DB 6; Length 955;  
Best Local Similarity 20.7%; Pred. No. 6.9;  
Matches 56; Conservative 44; Mismatches 72; Indels 99; Gaps 14;

QY 62 LKHLIQNOEYETELIMCSNNOAHLENENFVTFDYDPDAHLGYDFKQKLEFNAYPKF 121  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-708

QY 122 HEIYNQRTSVYMCVAIAIYKKEI-YLSGID-FYQNGSSYAFDTKQENLLKLPDPK 179  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-708

QY 180 -----NDRSH-----YIGHSKNTDI-----KALEFLEKTYKIKLYCLCPNSLLANFIEL 223  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-708

QY 224 APNLSNFIIOE-----KNYTKDILIPSEAYGKFSKNIF--KKIKIKENVYVKL 273  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-708

QY 274 IKDLR-----LPSDIKHYPK 289  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-471-571A-708

RESULT 6  
US-11-293-697-4502  
; Sequence 4502, Application US/11293697  
; Publication No. US20060105376A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: Novel full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/11/293,697  
; CURRENT FILING DATE: 2005-12-05  
; PRIOR APPLICATION NUMBER: US/10/108,260  
; PRIOR FILING DATE: 2002-03-28  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1

QY 37 DKYLGKCKKAVFYTPNFFEQYYTLKHLIQNOEYETELIMCSNNOAHLE--NENF--- 91  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 51 DEIMLSKFLKNIF---NVFFKYNHSERRGOPEKLERLIQPSFTSDTHELELQDFDKA 107  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 92 -----VKTFDYDPDA-----HLGYDFKQKLEFNAYFKPHEIYFNQRTSVYVM 136  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 108 DKLDKRPILSPKLRVLEELSESEVHLKSELKSIQIH-----YIVERLESCHI 157  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 137 CAVAIALGYKKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPKNDRSHYIGHSKNTDICA 196  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 158 TKEDLPKIYQNLVLMNEKAEQKGN-SFOGKYSETVKEIMSVNNFNHFF-IDKHLEIKL 215  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 197 LEFLEKTYKIKLYCLCPNSLLANFIE-----LAPNLSNFIIOEKNNYTKDILI 245  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 216 RSFLKE-----ILQNYELKNISESLFNSETATSIYPNISS-----LRTKSVSI 259  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 246 PSSEAYGKFSKNINFKKIKIKENVYVKLIK-----DLLRLPSDI-KH 286  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 260 SFHELEQDISKSGFRRFEI--NMKYPLSKSLQNYLIALSENELLHLKADLSKH 311  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

RESULT 7  
US-11-056-355B-91714  
; Sequence 91714, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence determined DNA fragments and corresponding  
; TITLE OF INVENTION: Polypeptides encoded thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 91714  
; LENGTH: 512  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(512)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728960  
US-11-056-355B-91714

Query Match 5.6%; Score 88; DB 7; Length 512;  
Best Local Similarity 21.9%; Pred. No. 7.5;  
Matches 77; Conservative 49; Mismatches 125; Indels 100; Gaps 19;

QY 5 IIAAGN-GPSLKEIDVSR-----PND-----FDVFCNQFYFEDKYVLGKCKKAVFYTP 52  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 114 LIAGRVDEALKQLEMMVRKLNPNNEATIRTFVHGIFRCL-----PPCKA-FEVL-161  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 53 NFFFEQYYTLKHLIQNOEYETELIMCSNNOAHLENENFVTFDY--YFDDAHLGYDFPK 110  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 162 VGFMEKDSNL---ORVGYDAVLYCLSNNSMAK-ETGQFLRKIGERYIPDS----- 208  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 111 QLKFNFA-----YFKPHEIYFNQRTSVG-----YMCVAIAL-----GYKEI 148  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 209 --STFNAMSCILLKGHDIVETCRIFDGFVSRGVKPGFNGYLVVQALLNAQRFSEGDRYL 266  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502

QY 149 YLSGIDFYQNGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSKH 190  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-293-697-4502



Db 267 KQMGVDGLLS-SVYSYNAVIDCLCKARRIENAMFLTEMQDRGSPNLVTFNFIUSG8EV 325  
Qy 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKN 238  
Db 326 RGDVKVHVGLLEKLLVHGFKPDVITFSLIINCLCKRAKEIKDAFCFKEMLEWGI--EPNE 383  
Qy 239 YTKDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHYPK 289  
Db 384 ITYNILIRSCSTGTDTRSVKL-FAKMKEN---GLSPDLAYAYNATIOQSFK 430

RESULT 8  
US-11-056-355B-95470  
; Sequence 95470, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; FILE REFERENCE: Polypeptides Encoded Thereby  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 95470  
; LENGTH: 512  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(512)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728960  
US-11-056-355B-95470

Query Match 5.6%; Score 88; DB 7; Length 512;  
Best Local Similarity 21.9%; Pred.No. 7.5;  
Matches 77; Conservative 49; Mismatches 125; Indels 100; Gaps 19;  
Qy 5 IIAGN-GPSLKEIDYSRL-----PND-----FDVFRNQFYFEDKYVLGKKCAVFTTP 52  
Db 114 LIAGRVDEALKOLEMRVKLPNEATIRTFVHGIFRCL-----PPCKA-FEVL 161  
Qy 53 NPFPEQYTYLKHLIQNOEYETELIMCSYNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110  
Db 162 VGFMEKDSNL-----QRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERYIPDS----- 208  
Qy 111 QLKEFNA-----YFKPHEIYFNQRTSGV-----YMCVAIAIAL-----GYKEI 148  
Db 209 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVQALLNAQRFSEGDRL 266  
Qy 149 YLSGIDFYQNGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190  
Db 267 KQMGVDGLLS-SVYSYNAVIDCLCKARRIENAMFLTEMQDRGSPNLVTFNFIUSG8EV 325  
Qy 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKN 238  
Db 326 RGDVKVHVGLLEKLLVHGFKPDVITFSLIINCLCKRAKEIKDAFCFKEMLEWGI--EPNE 383  
Qy 239 YTKDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHYPK 289  
Db 384 ITYNILIRSCSTGTDTRSVKL-FAKMKEN---GLSPDLAYAYNATIOQSFK 430

RESULT 9  
US-11-056-355B-91713  
; Sequence 91713, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai

; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; FILE REFERENCE: Polypeptides Encoded Thereby  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 91713  
; LENGTH: 530  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(530)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728959  
US-11-056-355B-91713

Query Match 5.6%; Score 88; DB 7; Length 530;  
Best Local Similarity 21.9%; Pred.No. 7.8;  
Matches 77; Conservative 49; Mismatches 125; Indels 100; Gaps 19;  
Qy 5 IIAGN-GPSLKEIDYSRL-----PND-----FDVFRNQFYFEDKYVLGKKCAVFTTP 52  
Db 132 LIAGRVDEALKOLEMRVKLPNEATIRTFVHGIFRCL-----PPCKA-FEVL 179  
Qy 53 NPFPEQYTYLKHLIQNOEYETELIMCSYNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110  
Db 180 VGFMEKDSNL-----QRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERYIPDS----- 226  
Qy 111 QLKEFNA-----YFKPHEIYFNQRTSGV-----YMCVAIAIAL-----GYKEI 148  
Db 227 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVQALLNAQRFSEGDRL 284  
Qy 149 YLSGIDFYQNGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190  
Db 285 KQMGVDGLLS-SVYSYNAVIDCLCKARRIENAMFLTEMQDRGSPNLVTFNFIUSG8EV 343  
Qy 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKN 238  
Db 344 RGDVKVHVGLLEKLLVHGFKPDVITFSLIINCLCKRAKEIKDAFCFKEMLEWGI--EPNE 401  
Qy 239 YTKDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHYPK 289  
Db 402 ITYNILIRSCSTGTDTRSVKL-FAKMKEN---GLSPDLAYAYNATIOQSFK 448

RESULT 10  
US-11-056-355B-95469  
; Sequence 95469, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; FILE REFERENCE: Polypeptides Encoded Thereby  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 95469  
; LENGTH: 530  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(530)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728959  
US-11-056-355B-95469

Query Match 5.6%; Score 88; DB 7; Length 530;

Best Local Similarity 21.9%; Pred. No. 7.8;  
Matches 77; Conservative 49; Mismatches 125; Indels 100; Gaps 19;  
QY 5 IIAAGN-GPSLKEIDYSLR-----PND-----FDVFRNQFYFEDKYILGKKCAVFTYP 52  
Db 132 LIAGRVDEALKQLEMMRVKLNPNENATIRTFVHGIFRCL-----PPCKA-FEVL 179  
QY 53 NFFFEQYVTLKHLIQNOEYETELIMCSYNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110  
Db 180 VGFMEKDSNL-----QRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERGYPDS----- 226  
QY 111 QLKEFNA-----YFKPHEIYFNQITSGV-----YMCVAIAIAL-----GYKEI 148  
Db 227 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVQALLNAQRFSEGDYVL 284  
QY 149 YLSGIDFYONGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190  
Db 285 KQMGVDGLLS-SVYSYNAVIDCLCKARRIENAAFLTEMQDRGISPNLVTFTFLSGYSV 343  
QY 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKNN 238  
Db 344 RGDVKKVHGVLKLLVHGFKPDVITFSLIINCLCRAKEIKDAFCFKEMLEWGI--EPNE 401  
QY 239 YTKDILIPSEAYGKFSKNINFKIKENYVYKLIKDLRLPSDIKHVFK 289  
Db 402 ITYNILIRSCCSTGDTDRSVKL-FAKMKN-----GLSPDLVAYNATIOQSFK 448

## RESULT 11

US-11-056-355B-91712  
; Sequence 91712, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 91712  
; LENGTH: 659  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(659)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728958  
US-11-056-355B-91712

Query Match 5.6%; Score 88; DB 7; Length 659;  
Best Local Similarity 21.9%; Pred. No. 10;  
Matches 77; Conservative 49; Mismatches 125; Indels 100; Gaps 19;  
QY 5 IIAAGN-GPSLKEIDYSLR-----PND-----FDVFRNQFYFEDKYILGKKCAVFTYP 52  
Db 261 LIAGRVDEALKQLEMMRVKLNPNENATIRTFVHGIFRCL-----PPCKA-FEVL 308  
QY 53 NFFFEQYVTLKHLIQNOEYETELIMCSYNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110  
Db 309 VGFMEKDSNL-----QRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERGYPDS----- 355  
QY 111 QLKEFNA-----YFKPHEIYFNQITSGV-----YMCVAIAIAL-----GYKEI 148  
Db 356 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVQALLNAQRFSEGDYVL 413  
QY 149 YLSGIDFYONGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190  
Db 414 KQMGVDGLLS-SVYSYNAVIDCLCKARRIENAAFLTEMQDRGISPNLVTFTFLSGYSV 472

QY 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKNN 238  
Db 473 RGDVKKVHGVLKLLVHGFKPDVITFSLIINCLCRAKEIKDAFCFKEMLEWGI--EPNE 530  
QY 239 YTKDILIPSEAYGKFSKNINFKIKENYVYKLIKDLRLPSDIKHVFK 289  
Db 531 ITYNILIRSCCSTGDTDRSVKL-FAKMKN-----GLSPDLVAYNATIOQSFK 577

## RESULT 12

US-11-056-355B-95468  
; Sequence 95468, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 95468  
; LENGTH: 659  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(659)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728958  
US-11-056-355B-95468

Query Match 5.6%; Score 88; DB 7; Length 659;  
Best Local Similarity 21.9%; Pred. No. 10;  
Matches 77; Conservative 49; Mismatches 125; Indels 100; Gaps 19;  
QY 5 IIAAGN-GPSLKEIDYSLR-----PND-----FDVFRNQFYFEDKYILGKKCAVFTYP 52  
Db 261 LIAGRVDEALKQLEMMRVKLNPNENATIRTFVHGIFRCL-----PPCKA-FEVL 308  
QY 53 NFFFEQYVTLKHLIQNOEYETELIMCSYNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110  
Db 309 VGFMEKDSNL-----QRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERGYPDS----- 355  
QY 111 QLKEFNA-----YFKPHEIYFNQITSGV-----YMCVAIAIAL-----GYKEI 148  
Db 356 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVQALLNAQRFSEGDYVL 413  
QY 149 YLSGIDFYONGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190  
Db 414 KQMGVDGLLS-SVYSYNAVIDCLCKARRIENAAFLTEMQDRGISPNLVTFTFLSGYSV 472

QY 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKNN 238  
Db 473 RGDVKKVHGVLKLLVHGFKPDVITFSLIINCLCRAKEIKDAFCFKEMLEWGI--EPNE 530  
QY 239 YTKDILIPSEAYGKFSKNINFKIKENYVYKLIKDLRLPSDIKHVFK 289  
Db 531 ITYNILIRSCCSTGDTDRSVKL-FAKMKN-----GLSPDLVAYNATIOQSFK 577

## RESULT 13

US-10-953-349-10851  
; Sequence 10851, Application US/10953349  
; Publication No. US20060107345A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: ENCODED THEREBY  
; FILE REFERENCE: 2750-1579PUS2  
; CURRENT APPLICATION NUMBER: US/10/953,349





GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:30:11 ; Search time 137.396 Seconds  
(without alignments)  
1959.153 Million cell updates/sec

Title: US-10-734-719-9  
Perfect score: 1563  
Sequence: 1 MKVLIAGNPSLKEIDYSR.....KLKLLRLPSDKHYFKGK 291

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.\*

1: uniprot\_prot.\*

2: uniprot\_tmbl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1563	100.0	291	2	Q919Q5_CAMJJE
2	1559	99.7	291	2	Q93D05_CAMJJE
3	1557	99.6	291	2	Q9F0M9_CAMJJE
4	1525	97.6	291	2	Q93M00_CAMJJE
5	1525	97.6	291	2	Q9LAK3_CAMJJE
6	1521	97.3	284	2	Q50FZ0_CAMJJE
7	1511	96.7	291	2	Q938X6_CAMJJE
8	1496	95.7	291	2	Q93CZ5_CAMJJE
9	1455.5	93.1	292	2	Q5DT12_CAMJJE
10	1192.5	76.3	263	2	Q2TKA8_CAMJJE
11	804	51.4	294	2	Q7BP25_CAMJJE
12	744.5	47.6	326	2	Q4QM36_HAB18
13	743	47.5	430	2	Q2RGF1_CAMJJE
14	740.5	47.4	430	2	Q32VR8_CAMJJE
15	730.5	46.7	320	2	Q4QNI8_HAB18
16	692.5	44.3	303	2	Q9CLP3_PASMU
17	485.5	31.1	231	1	Y352_HAEIN
18	169	10.8	44	2	Q4HGT1_CAMCO
19	152.5	9.8	1014	2	Q8RIY7_CAMJJE
20	152.5	9.8	1014	2	Q8RN21_CAMJJE
21	152.5	9.8	1176	2	Q51CW9_ENTHI
22	145.5	9.3	673	2	Q4HEM3_CAMCO
23	135	8.6	849	2	Q81SV7_PLAF7
24	133	8.5	271	2	Q3NR31_SHEP7
25	132.5	8.5	2755	2	Q81HP8_PLAF7
26	129.5	8.3	615	2	Q4XLN2_PLACH
27	129	8.3	417	2	Q9ALT0_CAMJJE
28	129	8.3	567	2	Q4XYM5_PLACH
29	129	8.3	1639	2	Q4YVD8_PLABE
30	129	8.3	2772	2	Q7RG21_PLAYO
31	128.5	8.2	1347	2	Q4Z0S9_PLABE

32	128	8.2	682	2	Q4YZJ1_PLABE	Q4YZJ1 plasmodium
33	128	8.2	819	2	Q6LF52_PLAF7	Q6LF52 plasmodium
34	128	8.2	842	2	Q815V4_PLAF7	Q815V4 plasmodium
35	127.5	8.2	560	2	Q81JN4_PLAF7	Q81JN4 plasmodium
36	127.5	8.2	804	2	Q54TP7_DICDI	Q54TP7 dictyosteli
37	127.5	8.2	919	2	Q4VZ88_PLABE	Q4VZ88 plasmodium
38	127	8.1	713	2	Q6BT89_DEBHA	Q6BT89 debaryomyce
39	126	8.1	551	2	Q81JG5_PLAF7	Q81JG5 plasmodium
40	125.5	8.0	515	2	Q65ZW7_BORGA	Q65ZW7 borrelia ga
41	125.5	8.0	601	2	Q891W3_CLOTE	Q891W3 clostridium
42	125.5	8.0	803	2	Q4YST0_PLABE	Q4YST0 plasmodium
43	125.5	8.0	859	2	Q54GX8_DICDI	Q54GX8 dictyosteli
44	125	8.0	885	2	Q81EM9_PLAF7	Q81EM9 plasmodium
45	125	8.0	1453	2	Q9G9H3_SCHCO	Q9G9H3 schizophylli

## ALIGNMENTS

RESULT 1  
Q919Q5\_CAMJJE  
ID Q919Q5\_CAMJJE PRELIMINARY; PRT; 291 AA.  
AC Q919Q5;  
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.  
DT 21-FEB-2006, sequence version 1.  
DE Alpha-2,3-sialyltransferase.  
GN Name=cst-II; Synonyms=cst, cstII;  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Campylobacteraceae; Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=ATCC 43446;  
RA MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;  
RA Gilbert M., Brisson J.-R., Karwaski M.-F., Michniewicz J.,  
Cunningham A.-M., Wu Y., Young N.M., Wakarchuk W.W.;  
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.  
RT Identification of the glycosyltransferase genes, enzymatic synthesis  
RT of model compounds, and characterization of nanomole amounts by 600-  
RT MHz (1)H and (13)C NMR analysis.";  
RT J. Biol. Chem. 275:3896-3906(2000).  
RL [2]  
RN NUCLEOTIDE SEQUENCE.  
RC STRAIN=ATCC 43446;  
RA Gilbert M.;  
RN Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=81-176;  
RA MEDLINE=21655167; PubMed=11796612; DOI=10.1128/IAI.70.2.787-793.2002;  
RA Guerry P., Szymanski C.M., Prendergast M.M., Hickey T.E., Ewing C.P.,  
Pattarini D.L., Moran A.P.;  
RT "Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide  
RT affects ganglioside mimicry and invasiveness in vitro.";  
RL Infect. Immun. 70:787-793(2002).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=GB11;  
RA PubMed=14742567; DOI=10.1128/IAI.72.2.1162-1165.2004;  
RA Gilbert M., Godschalk P.C., Karwaski M.F., Ang C.W., Van Belkum A.,  
Li J., Wakarchuk W.W., Endtz H.P.;  
RT "Evidence for acquisition of the lipooligosaccharide biosynthesis  
RT locus in Campylobacter jejuni GB11, a strain isolated from a patient  
RT with Guillain-Barre syndrome, by horizontal exchange.";  
RL Infect. Immun. 72:1162-1165(2004).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CF90-26;  
RA Koga M., Gilbert M., Takahashi M., Li J., Koike S., Hirata K.,  
Yuki N.;  
RT "Comprehensive Analysis of Bacterial Risk Factors for the Development

```
RT of Guillain-Barre Syndrome after Campylobacter jejuni Enteritis." ;
RL J. Infect. Dis. 193:547-555(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF167344; AAF34137.1; -; Genomic_DNA.
DR EMBL; AF305571; AAL09368.1; -; Genomic_DNA.
DR EMBL; AY421197; AAR82875.1; -; Genomic_DNA.
DR EMBL; AY661458; AAV80723.1; -; Genomic_DNA.
DR SMR; Q9L905; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34611 MW; 1BE4A04ACD728228 CRC64;

Query Match 100.08; Score 1563; DB 2; Length 291;
Best Local Similarity 100.08; Pred. No. 1.5e-104;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQYFEDKYVILGKCKKAVFYTPNFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQYFEDKYVILGKCKKAVFYTPNFFEQYY 60
QY 61 TLKHLIQNQYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120
DB 61 TLKHLIQNQYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120
QY 121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
DB 121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
QY 181 DRSHVIGHSKNTDIALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHVIGHSKNTDIALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNIINFKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
DB 241 KDILIPSSAYGKFSKNIINFKIKENYVYKLIKDLLRLPSDIKHYFKGK 291

RESULT 2
Q93D05 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93D05;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 13.
DE Alpha-2,3-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43456;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF215659; AAG43979.1; -; Genomic_DNA.
DR SMR; Q9F0M9; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34554 MW; E41B594ACD7280F8 CRC64;

Query Match 99.6%; Score 1557; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 4e-104;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQYFEDKYVILGKCKKAVFYTPNFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQYFEDKYVILGKCKKAVFYTPNFFEQYY 60
QY 61 TLKHLIQNQYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120
DB 61 TLKHLIQNQYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120
QY 121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
DB 121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
QY 181 DRSHVIGHSKNTDIALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHVIGHSKNTDIALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNIINFKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
DB 241 KDILIPSSAYGKFSKNIINFKIKENYVYKLIKDLLRLPSDIKHYFKGK 291

RESULT 3
Q9F0M9 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q9F0M9;
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 12.
DE Alpha-2,3-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43432;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF215659; AAG43979.1; -; Genomic_DNA.
DR SMR; Q9F0M9; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34554 MW; E41B594ACD7280F8 CRC64;

Query Match 99.6%; Score 1557; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 4e-104;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQYFEDKYVILGKCKKAVFYTPNFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQYFEDKYVILGKCKKAVFYTPNFFEQYY 60
QY 61 TLKHLIQNQYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120
DB 61 TLKHLIQNQYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120
QY 121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
DB 121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
QY 181 DRSHVIGHSKNTDIALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHVIGHSKNTDIALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNIINFKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
DB 241 KDILIPSSAYGKFSKNIINFKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
```

Db 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291

## RESULT 4

Q93MQ0 CAMJE PRELIMINARY; PRT; 291 AA.  
 AC Q93MQ0;  
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.  
 DT 01-DEC-2001, sequence version 1.  
 DT 07-FEB-2006, entry version 13.  
 DE Alpha-2,3-/alpha-2,8-sialyltransferase.  
 GN Name=catII;  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 43438;  
 RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;  
 RA Gilbert M., Karwaski M.-P., Bernatchez S., Young N.M., Taboada E.,  
 RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;  
 RT "The genetic basis for the variation in the lipo-oligosaccharide of  
 the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated  
 RT ganglioside mimics in the core oligosaccharide.";  
 RL J. Biol. Chem. 277:327-337(2002).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF400048; AAK91725.1; -; Genomic\_DNA.  
 DR SMR; Q93MQ0; 1-258.  
 DR GO; GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.  
 DR InterPro; IPR009251; CST-I.  
 DR Pfam; PF06002; CST-I; 1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 291 AA; 34573 MW; BD1EEDFEFF6F521E6 CRC64;

Query Match 97.6%; Score 1525; DB 2; Length 291;  
 Best Local Similarity 97.6%; Pred. No. 8.2e-102;  
 Matches 284; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYFEDKYLKCKKAVFYTPNFFEQY 60  
 Db 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYFEDKYLKCKKAVFYTPNFFEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFPKQKFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFPKQKFNAYFK 120  
 QY 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFN 180  
 Db 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291

## RESULT 5

Q9LAK3\_CAMJE

ID Q9LAK3\_CAMJE PRELIMINARY; PRT; 291 AA.  
 AC Q9LAK3;  
 DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.  
 DT 01-OCT-2000, sequence version 1.  
 DT 07-FEB-2006, entry version 15.  
 DE Alpha-2,3-/8-sialyltransferase (Alpha-2,3-/2,8-sialyltransferase).  
 GN Name=cst-II;  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=OH4384;  
 RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;  
 RA Gilbert M., Brissson J.-R., Karwaski M.-P., Michniewicz J.,  
 RA Cunningham A.-M., Wu Y., Young N.M., Wakarchuk W.W.;  
 RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.  
 RT Identification of the glycosyltransferase genes, enzymatic synthesis  
 RT of model compounds, and characterization of nanomole amounts by 600-  
 RT MHz (1)H and (13)C NMR analysis.";  
 RL J. Biol. Chem. 275:3896-3906(2000).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 700297;  
 RA Gilbert M., Michniewicz J., Wakarchuk W.W.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=CF 93-6;  
 RX Koga M., Gilbert M., Li J., Koike S., Takahashi M., Furukawa K.,  
 RA Hirata K., Yuki N.;  
 RT "Antecedent infections in Fisher Syndrome: a common pathogenesis of  
 RT molecular mimicry.";  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DR EMBL; AF130984; AAF31771.1; -; Genomic\_DNA.  
 DR EMBL; AF216647; AAL36462.1; -; Genomic\_DNA.  
 DR EMBL; AY644679; AAV52039.1; -; Genomic\_DNA.  
 DR PDB; 1R07; X-ray; A/B/C/D=1-259.  
 DR PDB; 1R08; X-ray; A/B=1-259.  
 DR GO; GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.  
 DR InterPro; IPR009251; CST-I.  
 DR Pfam; PF06002; CST-I; 1.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 291 AA; 34545 MW; AB0741D5FCEDAE6A CRC64;

Query Match 97.6%; Score 1525; DB 2; Length 291;  
 Best Local Similarity 97.3%; Pred. No. 8.2e-102;  
 Matches 283; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYFEDKYLKCKKAVFYTPNFFEQY 60  
 Db 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYFEDKYLKCKKAVFYTPNFFEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFPKQKFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFPKQKFNAYFK 120  
 QY 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFN 180  
 Db 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291

```
RESULT 6
Q50F20 CAMJE
ID Q50F20 CAMJE PRELIMINARY; PRT; 284 AA.
AC Q50F20
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Cj81-011 (Fragment).
DE Alpha-2,3-/alpha-2,8-sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=81-176;
RX PubMed=15872262; DOI=10.1128/JCM.43.5.2330-2338.2005;
RA Poly F., Thredgill D., Stintzi A.;
RT "Genomic Diversity in Campylobacter jejuni: Identification of C.
RT jejuni 81-176-Specific Genes.";
RL J. Clin. Microbiol. 43:2330-2338(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY681242; AAW56104.1; -; Genomic_DNA.
DR SMR; Q50F20; 1-258.
DR NON_TER 284
SQ SEQUENCE 284 AA; 33722 MW; 785BEA268C691E70 CRC64;

Query Match 97.3%; Score 1521; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.5e-101;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQFYFEDKYLGKKCAVFTPNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQFYFEDKYLGKKCAVFTPNPFPEQYY 60
QY 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFKQKKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFKQKKEFNAYFK 120
QY 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 284
DB 241 KDILIPSSAYGKFSKNINFKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 284

RESULT 7
Q938X6 CAMJE
ID Q938X6 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q938X6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Alpha-2,3-/alpha-2,8-sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43460;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQFYFEDKYLGKKCAVFTPNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQFYFEDKYLGKKCAVFTPNPFPEQYY 60
QY 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFKQKKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFKQKKEFNAYFK 120
QY 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 284
DB 241 KDILIPSSAYGKFSKNINFKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 284

RESULT 8
Q93CZ5 CAMJE
ID Q93CZ5 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93CZ5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Bifunctional alpha-2,3-/2,8-sialyltransferase.
OS Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF401529; AAL06004.1; -; Genomic_DNA.
DR SMR; Q93CZ5; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
```

```
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY044868; AAK96001.1; -; Genomic_DNA.
DR SMR; Q938X6; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34577 MW; 7BAE5F6021A56F08 CRC64;

Query Match 96.7%; Score 1511; DB 2; Length 291;
Best Local Similarity 95.5%; Pred. No. 8.3e-101;
Matches 278; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQFYFEDKYLGKKCAVFTPNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQFYFEDKYLGKKCAVFTPNPFPEQYY 60
QY 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFKQKKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFKQKKEFNAYFK 120
QY 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSAYGKFSKNINFKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
Q93CZ5 CAMJE
ID Q93CZ5 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93CZ5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Bifunctional alpha-2,3-/2,8-sialyltransferase.
OS Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF401529; AAL06004.1; -; Genomic_DNA.
DR SMR; Q93CZ5; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
```



```

DR InterPro: IPR009251; CST-I.
DR Pfam: PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34505 MW; 50162DBE8D81558F CRC64;

Query Match          95.7%; Score 1496; DB 2; Length 291;
Best Local Similarity 94.5%; Pred. No. 1e-99; Mismatches 8; Indels 0; Gaps 0;
Matches 275; Conservative

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTNPFEEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTNPFEEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNYYNOAHLENENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYYNOAHLENENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120

QY 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 239
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 239

QY 181 DSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 240
DB 181 DSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 240

QY 240 TKDILIPSSBAYGKFSKNINFKKIKENYVYKLIKDLLRLSPDIIKHYPFGK 291
DB 240 TKDILIPSSBAYGKFSKNINFKKIKENYVYKLIKDLLRLSPDIIKHYPFGK 291

QY 241 TKDILIPSSBAYGKFSKNINFKKIKENYVYKLIKDLLRLSPDIIKHYPFGK 292
DB 241 TKDILIPSSBAYGKFSKNINFKKIKENYVYKLIKDLLRLSPDIIKHYPFGK 292

RESULT 10
Q2TKA8 CAMJE PRELIMINARY; PRT; 263 AA.
AC Q2TKA8;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DE Putative alpha-2,3/8-sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GB5;
RA Gilbert M., Brochu D., Karwaski M.-F.;
RT "Sequencing of the Campylobacter jejuni GB5 LOS locus.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AY854153; AAX45339.1; -; Genomic DNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 263 AA; 31295 MW; 21E0289B1B9823E0 CRC64;

Query Match          76.3%; Score 1192.5; DB 2; Length 263;
Best Local Similarity 83.9%; Pred. No. 6.6e-78;
Matches 224; Conservative 14; Mismatches 16; Indels 13; Gaps 2;

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTNPFEEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTNPFEEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNYYNOAHLENENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120
DB 61 TLKHLIQNOEYETELIMCSNYYNOAHLENENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120

QY 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 240
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 240

QY 181 DSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 240
DB 181 DSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNY 240

QY 241 KOI-----LIPSSBAYGKFSKNINFKKI 263
DB 241 LKIYSLVLRME-----NFOKI 258

RESULT 11
Q7BP25 CAMJE PRELIMINARY; PRT; 294 AA.
AC Q7BP25; Q9PNF4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.

```







GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 14:13:03 ; Search time 95.0051 Seconds  
(without alignments)  
1418.823 Million cell updates/sec

Title: US-10-734-719-9  
Perfect score: 1563  
Sequence: 1 MKKVIAGNGPSLKEIDYSR.....KLIKDLLRLPSDIKHFKGK 291

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA\_Main:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	291	3	US-09-816-028A-9 Sequence 9, Appli
2	1563	100.0	291	4	US-10-303-161-9 Sequence 9, Appli
3	1563	100.0	291	4	US-10-303-118-9 Sequence 9, Appli
4	1563	100.0	291	4	US-10-303-128-9 Sequence 9, Appli
5	1563	100.0	291	4	US-10-303-134-9 Sequence 9, Appli
6	1563	100.0	291	4	US-10-303-162-9 Sequence 9, Appli
7	1563	100.0	291	4	US-10-820-536-9 Sequence 9, Appli
8	1563	100.0	291	4	US-10-845-408-9 Sequence 9, Appli
9	1563	100.0	291	4	US-10-845-412-9 Sequence 9, Appli
10	1563	100.0	291	5	US-10-846-219-9 Sequence 9, Appli
11	1563	100.0	291	5	US-10-821-604-9 Sequence 9, Appli
12	1563	100.0	291	5	US-10-847-983-9 Sequence 9, Appli
13	1563	100.0	291	5	US-10-821-573-9 Sequence 9, Appli
14	1563	100.0	291	5	US-10-850-807-9 Sequence 9, Appli
15	1563	100.0	291	5	US-10-850-125-9 Sequence 9, Appli
16	1563	100.0	291	5	US-10-830-825-9 Sequence 9, Appli
17	1563	100.0	291	5	US-10-962-334-9 Sequence 9, Appli
18	1563	100.0	291	5	US-10-830-997-9 Sequence 9, Appli
19	1563	100.0	291	5	US-10-962-235-9 Sequence 9, Appli
20	1563	100.0	291	5	US-10-961-882-9 Sequence 9, Appli
21	1559	99.7	291	3	US-09-816-028A-14 Sequence 14, Appl
22	1559	99.7	291	4	US-10-303-161-14 Sequence 14, Appl
23	1559	99.7	291	4	US-10-303-118-14 Sequence 14, Appl
24	1559	99.7	291	4	US-10-303-128-14 Sequence 14, Appl
25	1559	99.7	291	4	US-10-303-134-14 Sequence 14, Appl
26	1559	99.7	291	4	US-10-303-162-14 Sequence 14, Appl
27	1559	99.7	291	4	US-10-820-536-14 Sequence 14, Appl

28	1559	99.7	291	4	US-10-845-408-14	Sequence 14, Appl
29	1559	99.7	291	4	US-10-845-412-14	Sequence 14, Appl
30	1559	99.7	291	5	US-10-846-219-14	Sequence 14, Appl
31	1559	99.7	291	5	US-10-821-604-14	Sequence 14, Appl
32	1559	99.7	291	5	US-10-847-983-14	Sequence 14, Appl
33	1559	99.7	291	5	US-10-821-573-14	Sequence 14, Appl
34	1559	99.7	291	5	US-10-850-807-14	Sequence 14, Appl
35	1559	99.7	291	5	US-10-850-125-14	Sequence 14, Appl
36	1559	99.7	291	5	US-10-830-825-14	Sequence 14, Appl
37	1559	99.7	291	5	US-10-962-334-14	Sequence 14, Appl
38	1559	99.7	291	5	US-10-830-997-14	Sequence 14, Appl
39	1559	99.7	291	5	US-10-962-235-14	Sequence 14, Appl
40	1559	99.7	291	5	US-10-961-882-14	Sequence 14, Appl
41	1557	99.6	291	3	US-09-816-028A-12	Sequence 12, Appl
42	1557	99.6	291	4	US-10-303-161-12	Sequence 12, Appl
43	1557	99.6	291	4	US-10-303-118-12	Sequence 12, Appl
44	1557	99.6	291	4	US-10-303-128-12	Sequence 12, Appl
45	1557	99.6	291	4	US-10-303-134-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-816-028A-9  
; Sequence 9, Application US/09816028A  
; Patent No. US20020042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-09-816-028A-9

Query Match 100.0%; Score 1563; DB 3; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3, 2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKKVIAGNGPSLKEIDYSLPNDPFRNCQFEDKYLGGKCKAVFTPNPFEEQY	60
Db	1	MKKVIAGNGPSLKEIDYSLPNDPFRNCQFEDKYLGGKCKAVFTPNPFEEQY	60
QY	61	TLKHLIQNEQYETELIMCSNQAHLNENFVKTFDYFPDAHLGYDFPKQKEFNAYFK	120
Db	61	TLKHLIQNEQYETELIMCSNQAHLNENFVKTFDYFPDAHLGYDFPKQKEFNAYFK	120
QY	121	FHEIYNQRTSGVYMCVAIALGYKIYLSGIDFYONGSSYAFDTKQENLLKLPDPFN	180
Db	121	FHEIYNQRTSGVYMCVAIALGYKIYLSGIDFYONGSSYAFDTKQENLLKLPDPFN	180
QY	181	DRSHYGHGSKNTDIIKALEFEKTKYKLYCLCPNSLANFTELAPNLNSFIIOEKNNYT	240
Db	181	DRSHYGHGSKNTDIIKALEFEKTKYKLYCLCPNSLANFTELAPNLNSFIIOEKNNYT	240
QY	241	KDILIPSSSAYGKFSKNINFKIKENYVYKLIKDLLRLPSDIKHFKGK	291
Db	241	KDILIPSSSAYGKFSKNINFKIKENYVYKLIKDLLRLPSDIKHFKGK	291

Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 2

US-10-303-161-9  
; Sequence 9, Application US/10303161  
; Publication No. US20030148459A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,161  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-161-9

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKCKKAVFYTPNFFEQY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKCKKAVFYTPNFFEQY 60  
QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFFKOLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFFKOLKEFNAYFK 120  
QY 121 FHEIYNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 3

US-10-303-118-9  
; Sequence 9, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-118-9

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKCKKAVFYTPNFFEQY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKCKKAVFYTPNFFEQY 60  
QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFFKOLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFFKOLKEFNAYFK 120  
QY 121 FHEIYNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 4

US-10-303-128-9  
; Sequence 9, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-128-9

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYILGKKCAVFTPNFFPEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYILGKKCAVFTPNFFPEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYDPDAHLGYDFFKQKKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYDPDAHLGYDFFKQKKEFNAYFK 120  
QY 121 FHEIYNQRTITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYNQRTITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

## RESULT 5

US-10-303-134-9  
; Sequence 9, Application US/10303134  
; Publication No. US20030157657A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-134-9

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYILGKKCAVFTPNFFPEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYILGKKCAVFTPNFFPEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYDPDAHLGYDFFKQKKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYDPDAHLGYDFFKQKKEFNAYFK 120  
QY 121 FHEIYNQRTITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYNQRTITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291  
RESULT 6  
US-10-303-162-9  
; Sequence 9, Application US/10303162  
; Publication No. US20030157658A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-162-9

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYILGKKCAVFTPNFFPEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYILGKKCAVFTPNFFPEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYDPDAHLGYDFFKQKKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYDPDAHLGYDFFKQKKEFNAYFK 120  
QY 121 FHEIYNQRTITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYNQRTITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

## RESULT 7

US-10-820-536-9  
; Sequence 9, Application US/10820536  
; Publication No. US20040203103A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128

```

; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-820-536-9

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

QY 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
US-10-845-408-9
; Sequence 9, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-845-408-9

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

QY 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
US-10-845-408-9
; Sequence 9, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-845-408-9

```

```

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

QY 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 9
US-10-845-412-9
; Sequence 9, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-845-412-9

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPTNPFPEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFFKQKKEFNAYFK 120

QY 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 PHEIYFNQRTITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

```



QY 181 DRSHYGHGKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240  
 Db 181 DRSHYGHGKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 10  
 US-10-846-219-9  
 ; Sequence 9, Application US/10846219  
 ; Publication No. US20040219638A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/846,219  
 ; CURRENT FILING DATE: 2004-05-14  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-846-219-9

Query Match 100.0%; Score 1563; DB 5; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLKGKCAVYTNPFPEQY 60  
 Db 1 MKKVIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLKGKCAVYTNPFPEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFDYFPDAHLGYDFKQLKEFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFDYFPDAHLGYDFKQLKEFNAYFK 120  
 QY 121 FHEIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 Db 121 FHEIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYGHGKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240  
 Db 181 DRSHYGHGKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 11  
 US-10-821-604-9  
 ; Sequence 9, Application US/10821604  
 ; Publication No. US20040229263A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/821,604  
 ; CURRENT FILING DATE: 2004-04-08  
 ; PRIOR APPLICATION NUMBER: 10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-10-821-604-9  
 Query Match 100.0%; Score 1563; DB 5; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLKGKCAVYTNPFPEQY 60  
 Db 1 MKKVIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLKGKCAVYTNPFPEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFDYFPDAHLGYDFKQLKEFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFDYFPDAHLGYDFKQLKEFNAYFK 120  
 QY 121 FHEIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 Db 121 FHEIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYGHGKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240  
 Db 181 DRSHYGHGKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 12  
 US-10-847-983-9  
 ; Sequence 9, Application US/10847983  
 ; Publication No. US20040229272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/847,983  
 ; CURRENT FILING DATE: 2004-05-17  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni

```

; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-847-983-9

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLYGKCKAVFYTPNFFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLYGKCKAVFYTPNFFFEQYY 60

QY 61 TLKHLIQNQYETELIMCSNTNQAHLENENFVKTFFDYPPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQNQYETELIMCSNTNQAHLENENFVKTFFDYPPDAHLGYDFFKQKKEFNAYFK 120

QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 13
US-10-821-573-9
; Sequence 9, Application US/10821573
; Publication No. US20040229313A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRP
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-821-573-9

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLYGKCKAVFYTPNFFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLYGKCKAVFYTPNFFFEQYY 60

QY 61 TLKHLIQNQYETELIMCSNTNQAHLENENFVKTFFDYPPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQNQYETELIMCSNTNQAHLENENFVKTFFDYPPDAHLGYDFFKQKKEFNAYFK 120

QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 14
US-10-850-807-9
; Sequence 9, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; PRIOR FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRP
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-850-807-9

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 3.2e-126;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLYGKCKAVFYTPNFFFEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLYGKCKAVFYTPNFFFEQYY 60

QY 61 TLKHLIQNQYETELIMCSNTNQAHLENENFVKTFFDYPPDAHLGYDFFKQKKEFNAYFK 120
DB 61 TLKHLIQNQYETELIMCSNTNQAHLENENFVKTFFDYPPDAHLGYDFFKQKKEFNAYFK 120

QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIYLCPCNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 15
US-10-850-125-9
; Sequence 9, Application US/10850125
; Publication No. US20040259203A1

```

GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,125  
; CURRENT FILING DATE: 2004-05-19  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-850-125-9

Query Match 100.0%; Score 1563; DB 5; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.2e-126;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKCKKAVFYTNPFPEQYY 60  
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKCKKAVFYTNPFPEQYY 60  
QY 61 TLKHLIQNOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDPFKQLKEFNAYFK 120  
DB 61 TLKHLIQNOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDPFKQLKEFNAYFK 120  
QY 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQSNLLKLAPDPKN 180  
DB 121 FHEIFYNQRITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQSNLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNYT 240  
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKENVYKLIKDLLRLPSDIKHYPKGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKENVYKLIKDLLRLPSDIKHYPKGK 291

Search completed: July 19, 2006, 14:20:07  
Job time : 96.0051 secs



; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: E64149  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-231 <TIGR>  
A:Cross-references: UNIPROT:P24324; UNIPARC:UPI0000139681; GB:U32720; GB:L42023; NID:g15  
A:Experimental source: strain Rd KW20  
R:Maskell, D.J.; Szabo, M.J.; Butler, P.D.; Williams, A.E.; Moxon, E.R. Mol. Microbiol. 5, 1013-1022, 1991  
A:Title: Molecular analysis of a complex locus from Haemophilus influenzae involved in pathogenesis  
A:Reference number: S15287; MUID:92065797; PMID:1956282  
A:Accession: S15287  
A:Molecule type: DNA  
A:Residues: 1-231 <TIGR>  
A:Cross-references: UNIPARC:UPI000016782; EMBL:X57315; NID:g43587; PIDN:CAA40567.1; PID  
A:Experimental source: strain RM7004  
C:Superfamily: Haemophilus influenzae hypothetical protein HI0352

Query Match 31.1%; Score 485.5; DB 2; Length 231;  
Best Local Similarity 43.7%; Pred. No. 8.5e-28;  
Matches 101; Conservative 47; Mismatches 74; Indels 9; Gaps 6;  
65 LIQNOEYE-TLIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFKOLKEFNAYFKPHE 123  
Db 3 LIKNEVEYADIIUSSFVNLGDSLEK-IKVVQKLLTVDIGHYVNLKLPDAPDAVLAQNE 61  
124 IYFNQRTITSGVYMCVAIAIAGYKEIYLSGIDFYQ-NGSSYAFDTKQENLLKLAPDFKND 182  
Db 62 LYENKRITSGVYMCVAIAIAGYKEIYLSGIDFYQ-NGSSYAFDTKQENLLKLAPDFKND 121  
183 SHYIGHSKNTDIKALEFLEKTYKI-KLVCLCPNSLLANFIEIAPDNINFTKSE 240  
Db 122 SQSDIHSMEYDLNLYFLQKHGYNIIYICSPESPLCNFYPLSP-LNNPITFIEKKNYT 180  
241 KDILIPSSSEAYGKFSKNINPKIKIKENYVYKLIKDLRLPSDIKHFKGK 291  
Db 181 QDILIPPFVYKIG---IYKPRYIYQNLIFRLIIDLRLPNDIKHALKS 228

RESULT 3  
D90124  
hypothetical protein rad3 [imported] - Guillardia theta nucleomorph  
C:Species: nucleomorph Guillardia theta  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: D90124  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reid, N. Nature 410, 1091-1096, 2001  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; MUID:11323671; PMID:11323671  
A:Accession: D90124  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-706 <DOU>  
A:Cross-references: UNIPROT:Q98894; UNIPARC:UPI000008CF5B; GB:AF083031; NID:g13794312; F  
C:Genetics:  
A:Map position: 3  
A:Gene: rad3  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 7.8%; Score 122.5; DB 2; Length 706;  
Best Local Similarity 22.0%; Pred. No. 0.31;  
Matches 75; Conservative 56; Mismatches 111; Indels 99; Gaps 17;  
17 DYSLPNDPFRNCNQYFEDKYGLGKCKAVFTPNPFQYQYTLKHLIQNE 69  
Db 203 DYNFLILKFDQI-----IHSSKNYNTLNKKKIF-----LLFSNFSFSAALKSKSIISLSL 254

QY 70 -----EYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFKOLKEFNAYF 119  
Db 255 KTIIDSYSRLIIVERDIKKKIFQINVLKTFNNF---HLKNFFYIKKKNKILNNLVLT 310  
QY 120 -----KF-HEIYFNQRTITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQEN 170  
Db 311 KLFNQKFLHLEY---LNSYRPFLETLKLIKSKKEKTISSINIYKD-----IFLDEN 361  
QY 171 LKLLAP-----DFKNDRSHVIGHSKNTDIKALEFLEKTYKI-----KLVCLCPNSLLANFIE 222  
Db 362 EIEIAPVYQNFNDNISHYCKYTSFFDNKSLNGIKKISNFIILTKFELNQNNFINIKSE 421  
QY 223 LAPN-----LNSNFIIOEKNNY--TKDILIPSSSE 249  
Db 422 LDKNNSEKFKLTLEIDFNFGIEFFPELFRSIFITINDFEISKANNFFNCKNNFLGSIK 481  
QY 250 AYGKFSKNINFKIKI-----KENYV---YKLIKDLRL 280  
Db 482 NYQK-NKNCNYILTKYVQNQNKASKKDYEDYDFIGSILKL 521

## RESULT 4

S72284  
DNA-directed RNA polymerase (SC 2.7.7.6) beta'-2 chain - Plasmodium falciparum plastid  
N:Alternate names: RNA polymerase rpoC2  
C:Species: Plasmodium falciparum  
C:Date: 14-Apr-1998 #sequence\_revision 24-Apr-1998 #text\_change 09-Jul-2004  
C:Accession: S72284  
R:Wilson, R.J.M.; Denny, P.W.; Preiser, P.R.; Rangachari, K.; Roberts, K.; Roy, A.; Why  
J. Mol. Biol. 261, 155-172, 1996  
A:Title: Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium  
A:Reference number: S72277; MUID:96346169; PMID:8757284  
A:Accession: S72284  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-960 <WIL>  
A:Cross-references: UNIPROT:Q25802; UNIPARC:UPI0000075E46; EMBL:X95275; NID:g1171583; P  
A:Note: biosynthesis of this protein involves a -1 frameshift in the codon for residue  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C:Genetics:  
A:Gene: rpoC2  
A:Genome: plastid  
A:Note: this apparently degenerate plastid is referred to as the apicoplast  
C:Superfamily: DNA-directed RNA polymerase beta chain  
C:Keywords: nucleotidyltransferase; plastid; transcription

Query Match 7.8%; Score 122.5; DB 2; Length 960;  
Best Local Similarity 20.4%; Pred. No. 0.44;  
Matches 71; Conservative 48; Mismatches 100; Indels 129; Gaps 17;  
QY 25 FDFRCNQYFEDKYY-----LGKCKAVFTPNFF--FQYQYTLKHLIQNE 71  
Db 458 YNIYTYLYYHKKFYNLKNGIILNNNNKYNVIYFLINFLNFSNY---YKINYNNY 514  
QY 72 ETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFKOLKEFNAYFKHEIYFNORIT 131  
Db 515 -----NFINSNY-YFKKMFILKNFNNIQLNKLIFYNNFIYKYE 555  
QY 132 SGVYMCVAIAIAGYKEIYLSGID-----FYONGSSYAFD-----TKQENLLKLAPDFKND 181  
Db 556 KKLF-----IYLNIIINIIKKYLNFKYKYTKNKLFFIKKYNFNFLYVEIFKYN 603  
QY 182 RSHY--IGHSKNTDIKALEFLEKTYK-----IKLYCLCPNSLLANFIELANLNSNFI--- 232  
Db 604 WYKYLILANKYNYLIYNNYIKYLYKNININLYFI--KNLFYN-----NNNFTHH 653  
QY 233 IOEKNK-----YTKDILIPSSSEAYGK-----SKNI 258  
Db 654 IYKNNYIYNNNNMLYQYNKILINLLYNNKLIFYNNINNNYNNLYNNDITGLQSIINI 713  
QY 259 NPKKIKIKENV-----YYKLIKDLRLPS-----DIKH 287

Db 714 IFENKNKIDNIFPISNNIVVIFIKYNNVNNLIYIYINICNKNYNHY 761

RESULT 5

G90127

hypothetical protein orf590 [imported] - Guillardia theta nucleomorph

C:Species: nucleomorph Guillardia theta

A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C:Accession: G90127

R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reid

Nature 410, 1091-1096, 2001

A:Title: The highly reduced genome of an enslaved algal nucleus.

A:Reference number: A9082; MUID:11323671; PMID:11323671

A:Accession: G90127

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-590 <DOU>

A:Cross-references: UNIPROT:Q98567; UNIPARC:UPI00000AB3FB; GB:AF083031; NID:g13794337; F

C:Genetics:

A:Gene: orf590

A:Map position: 3

A:Genome: nucleomorph

C:Keywords: nucleomorph

Query Match 7.8%; Score 121.5; DB 2; Length 590;

Best Local Similarity 23.2%; Pred. No. 0.29;

Matches 82; Conservative 53; Mismatches 119; Indels 99; Gaps 19;

QY 17 DYSRLPNDYF-RCNQF-YPEDKYLGKCKKAVYTPNFF-----FEQYVTLKHLIQ 67

Db 5 DYLNRNLVSVCNFFTFVDLKKIFNFILQLLNISNFFSEKILILIFQNKPHLIHYIN 64

QY 68 --NOEYETELIMCSYNOAHLE-NENFVK-----TFYDYF----- 99

Db 65 RINELOQE-----KENEINFEKNDNYKSKSFNLMISLDFIFNEENFGILFNSKSLI 119

QY 100 ---PDHLYGDFPFKOLFNFAYFKPH-----EYFNQRTSGVYMCVAIALGYK-- 146

Db 120 FILDNKYILSNFKIWLKNFLFKFNKILIIITTSFNSQFSLSYFQFCNFAI--GSKII 177

QY 147 --EYLSGI----DF--YONGSSYAPDTKQENLLKLAPDKNDRSHYGHSKNTDKALEF 199

Db 178 NFQFYLRKRIKNEFLISDALSVEFNKKGILILINFKFD-----INELINSIYH 229

QY 200 LEKTYKIKLYCLCPNSLLANFIELAPNLSNFIQEKNNYK-----DI 243

Db 230 ISKIKKKKIFFF---PIINFI--MPDFVLNKKIIRLNKRYIKVLICRETSYKNFSLIFDL 283

QY 244 LIPSEAYGKFSKNINFKKIKIKENYVYKLIK-DLLRLPS---DIKHVFKGK 291

Db 284 IVELFKSNEKFKKINXYTHENLLNTKKNFKFNLIITFKYTKWTLRNLKKG 336

RESULT 6

T28418

ORF MSV257 leucine rich repeat gene family protein - Melanoplus sanguinipes entomopoxvir

C:Species: Melanoplus sanguinipes entomopoxvirus

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T28418

R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.

J. Virol. 73, 533-552, 1999

A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.

A:Reference number: 220484; MUID:99102612; PMID:9847359

A:Accession: T28418

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-707 <AFO>

A:Cross-references: UNIPROT:Q9YV15; UNIPARC:UPI00000F54AF; EMBL:AF063866; NID:g4049647;

C:Genetics:

A:Note: MSV257

Query Match 7.4%; Score 116; DB 2; Length 707;

Best Local Similarity 19.1%; Pred. No. 0.9;

Matches 98; Conservative 46; Mismatches 126; Indels 244; Gaps 22;

QY 6 IAGNGPSLKEIDYSRLPN--DFDVFRCN--QFVPEDKY----- 39

Db 182 ISNNEQLSKKIDYLRPLNLDIFNCSSCNKISFNFLNFTKLQILDISNNENIWTYALPLP 241

QY 40 -YLKG-KCAVAVTPNFFFEQYVTLKHL-IQNG-----EYETELIMCSYNOAHLENF 91

Db 242 PFLIKVNCSCNITNSGDFLRYVDNLELDISNPDLEIEMPTKIKKLINISECYIKDNF 301

QY 92 VKTF-----YDYF----- 99

Db 302 LKGLANNLQELNISYPNYFPRKSIDIDNLSNTLIKLYLRKCNKNVSPFNLILOEL 361

QY 100 -----PD-----AHLGYDFPKOLKEFN-----AYFPKH 122

Db 362 VISENQVNNUPDEILISNLSCKIINIEFPKCLKELNLSKTYISNIKLPKSLIKLD 421

QY 123 EYFNQR-----ITSGVYMCVA-----IALGYKEIYLSGIDFYQN--- 158

Db 422 ISVNLKNDNCLKDLNKLKELDISGNWEICFSSLPKSIKIKCKNCNITNTDFLNLNL 481

QY 159 -----GSSYAPDTKQENLLKLAPDKNDRSHYGHSKNTDKALEFLEK--- 202

Db 482 QELDMSVNYVININKSRRIDNLPKLIKL--NINNCDAVISFKGLDNLEEMIRNCTCL 539

QY 203 -----TYKI-----KLYCLC-----PNSL----- 216

Db 540 IDYLPKSLIKLHGHTFKIKOLNKLTTLCPLGIIKINIEDLPKSLIELDSTCKYIINE 599

QY 217 -----LANFIELAPNLSNFIQ-----EKNNYTKOIL---IPSSEAYG 252

Db 600 DCLKNLYNLKELTMSINCNIKQTLPESLIKIKIKLCKYSYDDFIKIDYKKNLSNEYG 659

QY 253 KP-----SKNINPKKIKIKENYV-YKLIKOL 277

Db 660 KYISFLPEIKKNYNAPKIHLDINYNLKKIKSI 693

RESULT 7

F90547

hypothetical protein MYPV 2860 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004

C:Accession: F90547

R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul

A:Reference number: A9512; MUID:21267165; PMID:11353084

A:Accession: F90547

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-781 <KUR>

A:Cross-references: UNIPROT:Q98Q35; UNIPARC:UPI00000D45AF; GB:AL445566; PID:g14089700;

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPV 2860

A:Genetic code: SGC3

Query Match 7.4%; Score 116; DB 2; Length 781;

Best Local Similarity 21.6%; Pred. No. 1;

Matches 71; Conservative 50; Mismatches 109; Indels 98; Gaps 17;

QY 24 DFDVFRCNQFY-----FEDKYVYLGKK-----KAVFYTPN----- 53

Db 134 DLSLVQNFYTFTRDKFNKDTKLENFEFKVESAKKALGSWEKLEKLTFSPEEKFDLNS 193

QY 54 ---FPFEQYVTLKHLIQNQYETELIMCSYNOAHLENFVKTYFYDPPDAHLGYD---- 107

Db 194 YENFYESNYSIQSLVINNEYR-----NYDK--LYNFYFEKQ-KDYRNKVNIDQDKIEL 243

QY 108 ---FFKQLKEFNAYFKFHEIYFNQRTSGVYMCVAIALGYKEIY-----LSGIDFYQNGS 160

Db 244 ATIREQELYNKYHKSSDYKA-----KELYDNKKLLG-EFKK--- 282  
Qy 161 SYAFDTKQENLLKLAPDFKND--RSHYIGHSKNTD-----IKALEFLEK-TYKIKL 208  
Db 283 --SFAATKENAVNMKSYINEFFLEYKLANSKAWDTTSNEWSYKLIATVKNISKYLLK 340  
Qy 209 YCLCPNSLLANFIELAPNLNFIQKNNYTKOIL-----IPSSAYAGKFSKNINFKKI- 263  
Db 341 VSFFKNLLNTLLFFLKFNENRFLTEIQIKYKYNEDFIYKFNLPKT-----FSKGPEKLV 395  
Qy 264 KIKENVYKLIKLLRLPSDIKHVEK 291  
Db 396 EYDAYFKIDKYLQRAQEHKNVLYLKEK 423  
RESULT 8  
A05037  
hypothetical protein 2136 - liverwort (Marchantia polymorpha) chloroplast  
C:Species: chloroplast Marchantia polymorpha  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: S01591; A05037  
R:Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T.  
J. Mol. Biol. 203, 299-331, 1988  
A:Title: Structure and organization of Marchantia polymorpha chloroplast genome. II. Gen  
A:Reference number: S01567; MUID:89068686; PMID:2974085  
A:Accession: S01591  
A:Molecule type: DNA  
A:Residues: 1-2136 <UME>  
A:Cross-references: UNIPROT:P09975; UNIPARC:UPI000013A622; EMBL:X04465; NID:g11640; PIDN  
R:Ohnaka, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shi  
Nature 322, 572-574, 1986  
A:Title: Chloroplast gene organization deduced from complete sequence of liverwort March  
A:Reference number: A38014  
A:Contents: annotation; gene organization, sites, features  
C:Genetics:  
A:Genome: chloroplast  
C:Keywords: chloroplast  
Query Match 7.3%; Score 114.5; DB 2; Length 2136;  
Best Local Similarity 21.1%; Pred. No. 4.1;  
Matches 74; Conservative 51; Mismatches 95; Indels 131; Gaps 15;  
Qy 15 EIDYSLRPNDFVFCNQFYFEDKYLYGKKCKAVFTNFPPFQYVTLKHLIQNOYEYETE 74  
Db 829 ELFISEINNDP---LMRFKKLYLYRYIKKEILFNP-----IENRQLLNFFEXTK 877  
Qy 75 LIMCSNVNQAHLNENFVKTYDYFPDAHLGYDFPKQKFNAYKPKFHEIYFNQRIYSGV 134  
Db 878 IL-----TFIDFLQDPELNY-----NNRFIFHLEKTKTKNNML 911  
Qy 135 YMCVAIAIAGYKEIYL-----SGIDFYONGSSYAFDT 166  
Db 912 YLRLLKFLKDKRNFLLINEIKSFIEKQNLFIKSQLSVLLVKNYSYKFFDNIFHFLEK 971  
Qy 167 KOENLLKLAPDKNDRSHYIGHS--KNTDIKAL-----EFLEKTYKIKLYCLCPNSL--- 216  
Db 972 QKERNIEILLNNQN---YFEKSLKKTLYLKNLNNNSYKFSYKIFIFQLL-NILNKN 1026  
Qy 217 -----LANFIELAPNLNFIQ---EKNNYTKDILIPSSAYAGKFSKNINFKKIKIKE 267  
Db 1027 YKTPQWISLIFYSKNL--NYKIQNKIEKNYC-----YNNKISYKKKIKTK 1071  
Qy 268 -----NVY--YKLIKOLLRLPSDIKHVEK 291  
Db 1072 VNFFKKNLFTQNNSWFFLEWWEYNTYILLQIQTETFFQITDVLVEFKK 1122  
RESULT 9  
T28423  
ORF MSV261 leucine rich repeat gene family protein - Melanoplus sanguinipes entomopoxvir  
C:Species: Melanoplus sanguinipes entomopoxvirus  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T28423  
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oms, E.; Kutish, G.F.; Rock, D.L.  
J. Virol. 73, 533-552, 1999  
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
A:Reference number: Z20484; MUID:99102612; PMID:9847359  
C:Accession: T28423  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-624 <AFO>  
A:Cross-references: UNIPROT:Q9YV11; UNIPARC:UPI000008EFC79; EMBL:AF063866; NID:g4049647;  
C:Genetics:  
A:Note: MSV261  
Query Match 7.3%; Score 114; DB 2; Length 624;  
Best Local Similarity 22.3%; Pred. No. 1.1;  
Matches 69; Conservative 37; Mismatches 118; Indels 86; Gaps 15;  
Qy 9 NCPSLKEIDYSR-----LPNDFVFCNQFYFEDKYLYGKKCKAVFTNFPPFQY 60  
Db 343 NLNLDKELDISRNFNINISLPSKSLITLINSYIKH-----TNFINDLY 387  
Qy 61 TLKHLIQNOYEYETE-----IMCSNQAHLNENFVKTYDYFPDAHLGYDFPKQK 114  
Db 388 NUKELDISRNVNINISLPSKSLIKLINSICDINKNTNFINDL-----DNLKELDI 436  
Qy 115 FNAYKPKFHEIYFNQRIYSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTTKQENLLKL 174  
Db 437 SNNY-----NINISS---LPKTLTTLNISCNKNNTNFIN---LDNLKE--LDI 478  
Qy 175 ADPFK-----NDRSHYIGHSKNT-----DIKALEFLEKTYKIKLYCLCPNSLL 217  
Db 479 SKNFNINISLPSKSLITLINSYINHTNFINDLYNLKELD-INSNNYINISL-PNTLI 536  
Qy 218 ANFIELAPNLNFIQKNNYTKDILIPSSAYAGKFSKNI---NFKKIKIKENV---Y 270  
Db 537 ELNISSCYKNTNFINDLN--LKKLEISKSLYDMILNHTFNTYKLLDKRNTSTFTY 594  
Qy 271 YKLIKDLLRL 280  
Db 595 YKIKSKFVKI 604  
RESULT 10  
H82926  
conserved hypothetical UUI43 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A:Reference number: A82870  
C:Accession: H82926  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1272 <GLA>  
A:Cross-references: UNIPARC:UPI00000C1BB3; GB:AE002115; GB:AF222894; NID:g6899102; PIDN  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UUI43  
A:Genetic code: SGC3  
Query Match 7.3%; Score 113.5; DB 2; Length 1272;  
Best Local Similarity 24.9%; Pred. No. 2.7;  
Matches 73; Conservative 33; Mismatches 88; Indels 99; Gaps 18;  
Qy 60 YTLKHLIQNOYEYETE-----MCSNQAHLNENFVKTYDY----- 98  
Db 373 FKIKHINIDSSNKKAIIELDSSSDNTKLLANNKQLIKSYDYNPNWSKIVSYEKDN 432  
Qy 99 ---FPDAHLGYDFPKQKFNAYKPKFHEIYFNQRIYSGVYMCVAIAIAGYKEIYLSGIDF 155  
Db 433 NKMIFDLH---DFPKDLKT---FIITHIRFDNITS-----LG--KIKENSFEY 473

Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: E64576  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1021 <TOM>  
A:Cross-references: UNIPROT:O25200; UNIPARC:UPI00000C0843; GB:AE000560; GB:AE000511; NCBI:1000000000  
C:Superfamily: *Helicobacter pylori* hypothetical protein HP0453

Query Match 7.2%; Score 112; DB 2; Length 1021;  
Best Local Similarity 19.6%; Pred. No. 2.7;  
Matches 60; Conservative 51; Mismatches 97; Indels 98; Gaps 13;

12 SLKEIDYSLPNDFVRCNQYFEDKYYLGK---KCKAVFYTPNFFEQYITLKHQ 67

16	IDYSLPN---	DPDVRNCQNFYEDKYLKCKKCAVY-----	TPNFFBQY	59
246	IDYHREDNFKKDILDKKN---	YEDLYYDEKLNSSYQQGLDKINKYVEVNNFNF--F	300	
60	YTLKHLINQBYETELIMCS-----	NYNOAHLENENFKTFY	96	
301	NTKKPLFKQFSQMFLEKDKLGLFKRYVLRNIDQKEVRARDFHFNDEKKQKNYFRSF-	359		
97	DYFPDAHLYGDOFFKOLKEFNAYFKPEHYFNORITSGVVMCAVALGKYETLYLSGDIFY	156		

Query Match 7.1%; Score 111.5; DB 2; Length 621;



Db 360 ----DKELNY--FQYIDHNDLIDINKISKEAYITQKIKNLLDAFNNNDVQITKYLQ 413  
Qy 157 QNGSSYAPDTKQENLLKLAPDFKDRSHVIGHSK-----NTDIIKALEFLEKTYKIKLYCL 211  
Db 414 QNDKEY-----QNKLIKQIIEFSN-----YFRNQKTPRDSNEYKQKLNDEKIFSQWY-- 462  
Qy 212 CPNSLLANFIELAPNLSNFIQEK--NNYTKDILIPSSAYGKFSKNINFKKIKIKENVY 270  
Db 463 ---FILMNLKHFELSFNWFTLPQINPITKEMLGPSB---KYLKRL--KKLKPYEDY 513  
Qy 271 Y 271  
Db 514 F 514

## RESULT 14

C70168  
DNA polymerase I (polA) homolog - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 09-Jul-2004  
C:Accession: C70168  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943; PMID:9403685  
A:Accession: C70168  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-908 <KLE>  
A:Cross-references: UNIPROT:O51498; UNIPARC:UPI000005750B; GB:AE001156; GB:AE000783; NID  
A:Experimental source: strain B31  
C:Superfamily: DNA-directed DNA polymerase I

Query Match 7.1%; Score 111.5; DB 2; Length 908;  
Best Local Similarity 26.1%; Pred. No. 2.5;  
Matches 76; Conservative 33; Mismatches 113; Indels 69; Gaps 16;  
Qy 24 DFDVFRNCQF-----YPED-----KYLGGKKKAVFTPNFFPEQYVTLKHLIQNOBYETEL 75  
Db 411 DYKILKNGGNFPPIPPYFDTMAAYLIIDTNSKV---SLDFLAEKY--LMH--KNIKYEDVI 463  
Qy 76 IMCSNYNOAHLENENFVKTFDYDPPDAHLGYDFF-----KOLKEFNAYKFEHY--FNQR 129  
Db 464 QKNDNFANISLE-----MATSYS-SEDAITFRLEFNIFTKUKEDKLDKLMHEIEMPFNKV 518  
Qy 130 I---TSGVVMCAVAI-----ALGYKEIYLSGIDFYQNGSSYAFDTKQENL-L 172  
Db 519 IIEEENGIVLDKEYKELGKELAEIENEIKSIGIDFNLSNPKQMEHILPEKLN 578  
Qy 173 KLAPDFKDRSHYIGHSKNNTDIKALEFI-----EKTYKIKLYCLCPNSLLANFIELAP 225  
Db 579 KLPEKMKD-----STDIKVLESRLREQHESLENLIKVRQIAKLAKLSTYTDNLIEL-- 627  
Qy 226 NLNSNFIQEKNNYTKDILIPSSAYGKF--SKNINFKKIKIKENVYKLIK 275  
Db 628 -----INYKTNRLHSTFIQTKTATGRITSINPNLQNIPIKDEKGRKIRK 671

## RESULT 15

D90129  
hypothetical protein orf1613 [imported] - Guillardia theta nucleomorph  
C:Species: nucleomorph Guillardia theta  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: D90129  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei  
Nature 410, 1091-1096, 2001  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; MUID:11323671; PMID:11323671

A:Accession: D90129  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1613 <DOU>  
A:Cross-references: UNIPROT:Q98S55; UNIPARC:UPI00000A244C; GB:AF083031; NID:gl3794350;  
C:Genetics:  
A:Gene: orf1613  
A:Map position: 3  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 7.1%; Score 111.5; DB 2; Length 1613;  
Best Local Similarity 21.8%; Pred. No. 4.9;  
Matches 77; Conservative 45; Mismatches 101; Indels 131; Gaps 19;

Qy 16 IDYSRLNDFDVFRCNQFYEDKYYLGGKC---KAVFYTPNFFPEQYVTLKHLIQNOEYE 72  
Db 1302 IKLNKIYDRFDKTIILHTIIYSVTNYFIKKCDKYMLMFYK-----FLMKRIFLDS--- 1350  
Qy 73 TELIMCSNYNOAHLENENF-----VKTFDYDPPDAHL-----GYDFFKQL 112  
Db 1351 -----VNINY-KISYEIFFQISVISFSNLFPPKTNLAKIMKNLFIKKFNLIILKEL 1403  
Qy 113 KEFNAYF---KPH-----EY-FNQRITSGVVMCAVALGAYKEIYLSGIDFYQN 158  
Db 1404 RNFSATFLILLKFKHLNLSINSLIKLKYNFYQKFNFIYLI-----PSKN 1446  
Qy 159 GSSYAFDTKQENLLKLAPDFKDRSHYIGHSKNTDIK-----ALEFLE----- 201  
Db 1447 GSSFI-----SLHLKXPLPENKYOKSKILS-SKYFTIKFFFEIYFLQSPFKFLELNFFS 1501  
Qy 202 -----KTY--KIKLYCLCP-NSLLANFIELAPNLSNFI-IOEKNNYTKD 242  
Db 1502 NFNLEDNKKRLLYFTKQYLLKLYSEIPTNLLIKNQINVLKXHLFEFISIQ----- 1554  
Qy 243 ILIPSSAYGKFSKNINFKKIKIKENVYKLIK-----DLLRLPSDIKHY 287  
Db 1555 --ILKKLFNMLNKFINFYRAKDYSNYSFKITKRKFIFNVYSLGFLNITKDIENF 1606

Search completed: July 19, 2006, 13:48:09  
Job time : 19.2031 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:30:08 ; Search time 107.972 Seconds  
(without alignments)  
1232.267 Million cell updates/sec

Title: US-10-734-719-9

Perfect score: 1563

Sequence: 1 MKKVIAGNGSLKEIDYSR.....KLIKDLLRLPSDIKHVFKGK 291

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_8.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*
- 9: Geneseqp2005s.\*
- 10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1563	100.0	291	3 AAY97212	Campyloba
2	1563	100.0	291	6 ABJ18482	Campyloba
3	1559	99.7	291	6 ABJ18485	Campyloba
4	1557	99.6	291	6 ABJ18484	Campyloba
5	1528	97.8	291	3 AAY97210	Campyloba
6	1528	97.8	291	6 ABJ18480	Campyloba
7	1525	97.6	291	3 AAY97204	Campyloba
8	1525	97.6	291	6 ABJ18479	Campyloba
9	1514	96.9	291	3 AAY97211	Campyloba
10	1514	96.9	291	6 ABJ18481	Campyloba
11	804	51.4	294	3 AAY97215	Campyloba
12	804	51.4	294	6 ABJ18483	Campyloba
13	743	47.5	430	2 AAY45221	Campyloba
14	692.5	44.3	303	6 ABR40195	alpha-2,3
15	119.5	7.6	958	8 ADN19559	Bacterial
16	113	7.2	2133	8 ADP25432	Plasmodiu
17	112	7.2	392	3 AAB15991	E. coli p
18	112	7.2	1021	2 AAW98555	H. pylori
19	111.5	7.1	908	6 ABU19261	Protein e
20	110	7.0	863	6 ABE08930	Allolococ
21	109.5	7.0	1121	3 AAB18241	Plasmodiu
22	109	7.0	1398	3 AAB18292	Plasmodiu
23	109	7.0	1817	3 AAB18301	Plasmodiu

24	108.5	6.9	422	4 AAU35736	Helicobac
25	108.5	6.9	422	6 ABR83248	H. pylori
26	108.5	6.9	422	6 ABR83247	H. pylori
27	108.5	6.9	422	6 ABU30795	Protein e
28	108.5	6.9	1284	6 ABU48910	Protein e
29	108	6.9	822	5 ABB77605	AmEPV sec
30	107.5	6.9	295	7 ADC96625	E. faeciu
31	107.5	6.9	388	8 ADK16435	Nanoarcha
32	107.5	6.9	400	6 ABU25530	Protein e
33	107	6.8	1802	3 AAB18217	Plasmodiu
34	107	6.8	1815	8 ADP25442	Plasmodiu
35	106.5	6.8	305	6 ABM69267	Photorhab
36	106.5	6.8	1294	7 ADC21264	Plasmodiu
37	106	6.8	2539	3 AAB18198	Plasmodiu
38	106	6.8	3973	3 AAB18253	Plasmodiu
39	105.5	6.7	615	6 ABU25054	Protein e
40	105.5	6.7	1346	6 AAB18236	Plasmodiu
41	105.5	6.7	1465	6 ABU19347	Protein e
42	105	6.7	1182	3 AAB18288	Plasmodiu
43	105	6.7	1817	3 AAB18255	Plasmodiu
44	105	6.7	2013	3 AAB18265	Plasmodiu
45	104.5	6.7	772	4 ABB57750	Drosophil

## ALIGNMENTS

### RESULT 1

AAU35736	AAU35736	standard; protein; 291 AA.
ABR83248	ABR83248	
ABR83247	ABR83247	
ABU30795	ABU30795	
ABU48910	ABU48910	
ABU77605	ABU77605	
ADK16435	ADK16435	
ADP25442	ADP25442	
ABM69267	ABM69267	
ADC21264	ADC21264	
AAB18198	AAB18198	
AAB18253	AAB18253	
ABU25054	ABU25054	
AAB18236	AAB18236	
ABU19347	ABU19347	
AAB18288	AAB18288	
AAB18255	AAB18255	
AAB18265	AAB18265	
ABB57750	ABB57750	

AAU35736	AAU35736	standard; protein; 291 AA.
ABR83248	ABR83248	
ABR83247	ABR83247	
ABU30795	ABU30795	
ABU48910	ABU48910	
ABU77605	ABU77605	
ADK16435	ADK16435	
ADP25442	ADP25442	
ABM69267	ABM69267	
ADC21264	ADC21264	
AAB18198	AAB18198	
AAB18253	AAB18253	
ABU25054	ABU25054	
AAB18236	AAB18236	
ABU19347	ABU19347	
AAB18288	AAB18288	
AAB18255	AAB18255	
AAB18265	AAB18265	
ABB57750	ABB57750	

Novel glycosyltransferase polypeptides and polynucleotides useful for biosynthesis of ganglioside and ganglioside mimics, as diagnostic reagents and as immunogen for producing antibodies.  
Disclosure; Page 96-97; 120pp; English.  
A reaction mixture for the synthesis of a sialylated oligosaccharide is useful for synthesizing sialylated oligosaccharides such as ganglioside, lysoganglioside or their mimics. Glycosyltransferases are useful for chemo-enzymatic synthesis of oligosaccharides, including gangliosides and other oligosaccharides that have biological activity. The enzymes and

CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyze the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 12-SEP-2003 to standardise OS field)

SQ Sequence 291 AA;

Query Match 100.0%; Score 1563; DB 3; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 9.9e-148;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRFCNQFYFEDKYILGKCKAVFYTPNFFFEQYY 60  
 DB 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRFCNQFYFEDKYILGKCKAVFYTPNFFFEQYY 60  
 QY 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFPKQKEFNAYFK 120  
 DB 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFPKQKEFNAYFK 120  
 QY 121 FHEIFYNQRTSGVTMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 DB 121 FHEIFYNQRTSGVTMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFKGK 291  
 DB 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFKGK 291

RESULT 2  
 ABJ18482

ID ABJ18482 standard; protein; 291 AA.

AC ABJ18482;

DT 07-FEB-2003 (first entry)

DE Campylobacter jejuni bifunctional sialtransferase cstII #4.

Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 GAINAC transferase; N-Acetylgalactosamine transferase;  
 galactosyltransferase; sialyltransferase; sialic acid synthase;  
 cytidine 5'-monophosphate sialic acid synthetase;  
 CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 ganglioside mimetics; inflammation; tumour metastasis.

OS Campylobacter jejuni.

FN WO200274942-A2.

PD 26-SEP-2002.

PF 22-FEB-2002; 2002WO-CA000229.

XX 21-MAR-2001; 2001US-00816028.

PA (CANA ) NAT RES COUNCIL CANADA.

PI Gilbert M, Wakarchuk WW;

DR WPI; 2003-040554/03.

DR N-PSDB; ABT13669.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.

PS Disclosure; Page 99; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GAINAC (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogens from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present amino acid sequence represents a Campylobacter  
 CC jejuni protein of the invention

SQ Sequence 291 AA;

Query Match 100.0%; Score 1563; DB 6; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 9.9e-148;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRFCNQFYFEDKYILGKCKAVFYTPNFFFEQYY 60  
 DB 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRFCNQFYFEDKYILGKCKAVFYTPNFFFEQYY 60  
 QY 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFPKQKEFNAYFK 120  
 DB 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDPDAHLGYDFPKQKEFNAYFK 120  
 QY 121 FHEIFYNQRTSGVTMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 DB 121 FHEIFYNQRTSGVTMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFKGK 291  
 DB 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFKGK 291

RESULT 3

ABJ18485

ID ABJ18485 standard; protein; 291 AA.

AC ABJ18485;

DT 07-FEB-2003 (first entry)

DE Campylobacter jejuni bifunctional sialtransferase cstII #6.

Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 GAINAC transferase; N-Acetylgalactosamine transferase;  
 galactosyltransferase; sialyltransferase; sialic acid synthase;  
 cytidine 5'-monophosphate sialic acid synthetase;  
 CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 ganglioside mimetics; inflammation; tumour metastasis.

OS Campylobacter jejuni.

XX WO200274942-A2.  
 PN 26-SEP-2002.  
 PD 22-FEB-2002; 2002WO-CA000229.  
 XX 21-MAR-2001; 2001US-00816028.  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 PA Gilbert M, Wakarchuk WW;  
 XX WPI; 2003-040554/03.  
 DR N-PSDB; ABT13671.  
 XX  
 PT New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Disclosure; Page 99; 107pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present amino acid sequence represents a Campylobacter  
 CC jejuni protein of the invention  
 XX  
 SQ Sequence 291 AA;  
 Query Match 99.7%; Score 1559; DB 6; Length 291;  
 Best Local Similarity 99.7%; Pred. No. 2.5e-147;  
 Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVFCNQFYEDKYLGGKCKAVFTNPFPEQY 60  
 Db 1 MKKVIIAGNGPSLKEIDYSLRPNDFVFCNQFYEDKYLGGKCKAVFTNPFPEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDPPDAHLGYDFPKQLKEFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDPPDAHLGYDFPKQLKEFNAYFK 120  
 QY 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
 Db 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGKG 291  
 Db 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGKG 291  
 RESULT 4  
 ABJ18484  
 ID ABJ18484 standard; protein; 291 AA.  
 XX  
 AC ABJ18484;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE Campylobacter jejuni bifunctional sialtransferase cstII #5.

XX Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAC transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 XX ganglioside mimetics; inflammation; tumour metastasis.  
 OS Campylobacter jejuni.  
 XX WO200274942-A2.  
 PN 26-SEP-2002.  
 PD 22-FEB-2002; 2002WO-CA000229.  
 XX 21-MAR-2001; 2001US-00816028.  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 PA Gilbert M, Wakarchuk WW;  
 XX WPI; 2003-040554/03.  
 DR N-PSDB; ABT13670.  
 XX  
 PT New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Disclosure; Page 99; 107pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present amino acid sequence represents a Campylobacter  
 CC jejuni protein of the invention  
 XX  
 SQ Sequence 291 AA;  
 Query Match 99.6%; Score 1557; DB 6; Length 291;  
 Best Local Similarity 99.7%; Pred. No. 3.9e-147;  
 Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MKKVIIAGNGPSLKEIDYSLRPNDFVFCNQFYEDKYLGGKCKAVFTNPFPEQY 60  
 Db 1 MKKVIIAGNGPSLKEIDYSLRPNDFVFCNQFYEDKYLGGKCKAVFTNPFPEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDPPDAHLGYDFPKQLKEFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDPPDAHLGYDFPKQLKEFNAYFK 120  
 QY 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
 Db 121 FHEIYFNORITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGKG 291  
 Db 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFFGKG 291

```

RESULT 5
AA97210
ID AA97210 standard; protein; 291 AA.
XX
AC AA97210;
XX
XX
DT 12-SEP-2003 (revised)
DT 22-DEC-2000 (first entry)
XX
XX Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.
XX
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside.
XX
XX Campylobacter jejuni; O:10 serotype.
XX
XX WO200046379-A1.
XX
XX 10-AUG-2000.
XX
XX 01-FEB-2000; 2000WO-CA000086.
XX
XX 01-FEB-1999; 99US-0118213P.
XX
XX 31-JAN-2000; 2000US-00495406.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
XX
XX WPI; 2000-524418/47.
XX
XX N-PSDB; AAA53724.
XX
XX Novel glycosyltransferase polypeptides and polynucleotides useful for
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
PT reagents and as immunogen for producing antibodies.
XX
XX Claim 13; Page 92-94; 120pp; English.
XX
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is
CC useful for synthesizing sialylated oligosaccharide such as ganglioside,
CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
CC other oligosaccharides that have biological activity. The enzymes and
CC nucleic acids that encode them are useful for studies of the pathogenesis
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
CC jejuni and the nucleic acids are used as probes to study expression of
CC genes involved in ganglioside mimetic synthesis. Antibodies raised
CC against the glycosyltransferases are also useful for analyzing the
CC expression patterns of these genes involved in pathogenesis. The nucleic
CC acids are also useful for designing antisense oligonucleotides for
CC inhibiting expression of the Campylobacter enzymes that are involved in
CC the biosynthesis of ganglioside mimics that can mask the pathogens from
CC the host's immune system. The oligosaccharides are useful as diagnosing
CC reagents or as therapeutics and as immunogens for producing antibodies.
CC Bacterial glycosyltransferase can be used to catalyze the formation of
CC oligosaccharides that are identical to the corresponding mammalian
CC structures and are easier and less expensive to produce in large
CC quantity, compared to the mammalian glycosyltransferase. The bacterial
CC origin of the enzymes facilitates expression of large quantities of the
CC enzymes using relatively inexpensive prokaryotic expression systems.
CC (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 291 AA;
SQ
Query Match 97.8%; Score 1528; DB 3; Length 291;
Best Local Similarity 97.9%; Pred. No. 3.2e-144;
Matches 285; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MKKVIIAGNGPSLKEIDYSRLPNDFDVFRCNQFYFEDKYYLGGKKKXAVFYNPGLFFEQYY 60
|||||

```

```

Db 1 MKKVIIAGNGPSLKEIDYSRLPNDFDVFRCNQFYFEDKYYLGGKKKXAVFYNPGLFFEQYY 60
QY 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYFDPDAHLGYDFFKQLKEFNAYFK 120
|||||
Db 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYFDPDAHLGYDFFKQLKEFNAYFK 120
QY 121 FHEIYNQRITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180
|||||
Db 121 FHEIYNQRITSGVYMCVAIAIAGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHVIGHGSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIQEKNNYT 240
|||||
Db 181 DRSHVIGHGSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIQEKNNYT 240
QY 241 KDILIPSEAYGKFSKNINFKIKIKENVVYKLIKDLLRPLPSDIKHVFKGK 291
|||||
Db 241 KDILIPSEAYGKFSKNINFKIKIKENVVYKLIKDLLRPLPSDIKHVFKGK 291

RESULT 6
ABJ18480
ID ABJ18480 standard; protein; 291 AA.
XX
AC ABJ18480;
XX
DT 07-FEB-2003 (first entry)
XX
DE Campylobacter jejuni bifunctional sialtransferase cstII #2.
XX
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
XX Campylobacter jejuni.
XX
XX WO200274942-A2.
XX
XX 26-SEP-2002.
XX
XX 22-FEB-2002; 2002WO-CA0000229.
XX
XX 21-MAR-2001; 2001US-00816028.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
XX
XX WPI; 2003-040554/03.
XX
XX N-PSDB; ABT13667.
XX
XX New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
XX Claim 5; Page 97; 107pp; English.
XX
XX The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present amino acid sequence represents a Campylobacter
CC jejuni protein of the invention

```

XX Sequence 291 AA;  
SQ  
Query Match 97.8%; Score 1528; DB 6; Length 291;  
Best Local Similarity 97.9%; Pred. No. 3.2e-144;  
Matches 285; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLGKKCAVFTPNFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLGKKCAVFTPNFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120  
QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLPDPKN 180  
DB 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 7  
AAAY97204  
ID AAAY97204 standard; protein; 291 AA.  
AC AAAY97204;  
DT 12-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE Campylobacter jejuni OH4384 CstII sialyltransferase.  
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside.  
XX  
OS Campylobacter jejuni; OH4384.  
XX  
PW WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CA000086.  
XX  
PR 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
PA (CAN ) NAT RES COUNCIL CANADA.  
XX  
PI Gilbert M, Wakarchuk WW;  
XX  
DR WPI; 2000-524418/47.  
DR N-PSDB; AAA53721, AAA53720.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
PS Claim 13; Page 91-92; 120pp; English.  
XX  
CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and

CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 291 AA;  
Query Match 97.6%; Score 1525; DB 3; Length 291;  
Best Local Similarity 97.3%; Pred. No. 6.3e-144;  
Matches 283; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLGKKCAVFTPNFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLGKKCAVFTPNFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDYPDAHLGYDFPKQLKEFNAYFK 120  
QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLPDPKN 180  
DB 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
RESULT 8  
ABJ18479  
ID ABJ18479 standard; protein; 291 AA.  
XX  
AC ABJ18479;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni bifunctional sialtransferase cstII #1.  
XX  
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.  
XX  
PW WO200274942-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 22-FEB-2002; 2002WO-CA000229.  
XX  
PR 21-MAR-2001; 2001US-00816028.

XX (CANA ) NAT RES COUNCIL CANADA.  
PA Gilbert M, Wakarchuk WW;  
PI WPI; 2003-040554/03.  
XX N-PSDB; A8T13666.  
XX  
XX New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
XX Claim 5; Page 96; 107pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
CC Acetyl-galactosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present amino acid sequence represents a Campylobacter  
XX jejuni protein of the invention  
XX  
XX Sequence 291 AA;

Query Match 97.6%; Score 1525; DB 6; Length 291;  
Best Local Similarity 97.3%; Pred. No. 6.3e-144;  
Matches 283; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFDVFCNQFYEDKYLGGKCKAVFYTPNFFEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDPFDVFCNQFYEDKYLGGKCKAVFYTPNFFEQYY 60  
QY 61 TLKHLIQNEQVETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFKQLKEFNAYFK 120  
Db 61 TLKHLIQNEQVETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFKQLKEFNAYFK 120  
QY 121 FHEIYFNQRTISGVVMCAVAIALGKYEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYFNQRTISGVVMCAVAIALGKYEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIIOEKNNYT 240  
Db 181 DSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINPKIKIKENYVYKLIKDLLRLPSDIKHYPKGK 291  
Db 241 KDILIPSEAYGKFSKNINPKIKIKENYVYKLIKDLLRLPSDIKHYPKGK 291

RESULT 9  
AAAY97211  
ID. AAAY97211 standard; protein; 291 AA.  
XX AC AAAY97211;  
XX  
XX 12-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
XX Campylobacter jejuni O:41 serotype alpha-2,3-sialyltransferase.  
XX  
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
XX immunity; immunogen; ganglioside.

OS Campylobacter jejuni; O:41 serotype.  
XX  
XX Key Location/Qualifiers  
FT CDS 1..876  
FT /\*tag= a  
FT /product= "alpha-2,3-sialyltransferase"  
XX  
XX WO200046379-A1.  
XX  
XX 10-AUG-2000.  
XX  
XX 01-FEB-2000; 2000WO-CA000086.  
XX  
XX 01-FEB-1999; 99US-0118213P.  
XX 31-JAN-2000; 2000US-00495406.  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX WPI; 2000-524418/47.  
XX DR N-PSDB; AA53725.  
XX  
XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
XX Claim 13; Page 94-95; 120pp; English.  
XX  
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesizing sialylated oligosaccharides such as ganglioside,  
CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 12-SEP-2003 to standardise OS field)  
XX  
XX Sequence 291 AA;

Query Match 96.9%; Score 1514; DB 3; Length 291;  
Best Local Similarity 95.9%; Pred. No. 8e-143;  
Matches 279; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFDVFCNQFYEDKYLGGKCKAVFYTPNFFEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDPFDVFCNQFYEDKYLGGKCKAVFYTPNFFEQYY 60  
QY 61 TLKHLIQNEQVETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFKQLKEFNAYFK 120  
Db 61 TLKHLIQNEQVETELIMCSNNOAHLENENFVKTFYDFPDALHGYDFKQLKEFNAYFK 120  
QY 121 FHEIYFNQRTISGVVMCAVAIALGKYEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYFNQRTISGVVMCAVAIALGKYEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIIOEKNNYT 240  
XX

Db 181 DNSHYGHGSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

Qy 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPGK 291

Db 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPGK 291

RESULT 10

ABJ18481

ID ABJ18481 standard; protein; 291 AA.

AC ABJ18481;

XX

XX

DT 07-FEB-2003 (first entry)

XX

DE Campylobacter jejuni bifunctional sialtransferase cstII #3.

XX

XX Enzyme; gene therapy; acyltransferase; glycosyltransferase;

KW GalNAc transferase; N-Acetylgalactosamine transferase;

KW galactosyltransferase; sialyltransferase; sialic acid synthase;

KW cytidine 5'-monophosphate sialic acid synthetase;

KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;

KW ganglioside mimetics; inflammation; tumour metastasis.

XX

OS Campylobacter jejuni.

XX

XX WO200274942-A2.

XX

XX 26-SEP-2002.

PD

XX

XX 22-FEB-2002; 2002WO-CA000229.

PF

XX

XX 21-MAR-2001; 2001US-00816028.

PR

XX (CANA ) NAT RES COUNCIL CANADA.

PA

XX Gilbert M, Wakarchuk WW;

PI

XX WPI; 2003-040554/03.

DR

XX N-PSDB; ABT13668.

DR

XX

XX New glycosyltransferases from Campylobacter, useful for synthesizing

PT gangliosides and ganglioside mimetics, and in studying the pathogenesis

PT mechanisms of organisms that synthesize ganglioside mimetics.

PT

XX Claim 5; Page 97-98; 107pp; English.

PS

XX The invention comprises the amino acid and coding sequences of

CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention

CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-

CC Acetyl-galactosamine) transferase; galactosyltransferase;

CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)

CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein

CC sequences of the invention are useful for ganglioside synthesis, studying

CC ganglioside mimetics, and for designing oligonucleotides to inhibit

CC expression of Campylobacter enzymes involved in the biosynthesis of

CC gangliosides. The C. jejuni oligosaccharides of the invention may be used as

CC diagnostic reagents (e.g. to locate areas of inflammation or tumour

CC metastasis). The present amino acid sequence represents a Campylobacter

CC jejuni protein of the invention

XX

SQ Sequence 291 AA;

Query Match 96.9%; Score 1514; DB 6; Length 291;

Best Local Similarity 95.9%; Pred. No. 8e-143;

Matches 279; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLKCKKCAVYTPNFFEQY 60

Db 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLKCKKCAVYTPNFFEQY 60

Qy 61 TLKHLIQNQYETELIMCSNQNQAHLENENFVKTFYDFPDHLYGDFFKQLKEFNAYFK 120

Db 61 TLKHLIQNQYETELIMCSNQNQAHLENENFVKTFYDFPDHLYGDFFKQLKEFNAYFK 120

Qy 121 FHEIYNQRIITSGVYMCVAIAALGYKIYLSGIDFYQNGSSYAFDTQENLLKLAPDFKN 180

Db 121 FHEIYNQRIITSGVYMCVAIAALGYKIYLSGIDFYQNGSSYAFDTQENLLKLAPDFKN 180

Qy 181 DRSHYIGHGSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

Db 181 DNSHYGHGSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

Qy 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPGK 291

Db 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPGK 291

RESULT 11

AA97215

ID AA97215 standard; protein; 294 AA.

XX

AC AA97215;

XX

DT 12-SEP-2003 (revised)

DT 22-DEC-2000 (first entry)

XX

DE Campylobacter jejuni NCTC11168 alpha-2,3-sialyltransferase.

XX

XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;

KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;

KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;

KW sialic acid synthase; CMP-sialic acid synthetase; mimetic; antibody;

KW immunity; immunogen; ganglioside.

XX

OS Campylobacter jejuni; NCTC11168.

XX

XX WO200046379-A1.

PN

XX

PD 10-AUG-2000.

XX

XX 01-FEB-2000; 2000WO-CA000086.

PF

XX 01-FEB-1999; 99US-0118213P.

PR

XX 31-JAN-2000; 2000US-00495406.

PR

XX (CANA ) NAT RES COUNCIL CANADA.

PA

XX Gilbert M, Wakarchuk WW;

PI

XX WPI; 2000-524418/47.

DR

XX Novel glycosyltransferase polypeptides and polynucleotides useful for

PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic

PT reagents and as immunogen for producing antibodies.

PT

XX Disclosure; Page 97-98; 120pp; English.

PS

XX A reaction mixture for the synthesis of a sialylated oligosaccharide is

CC useful for synthesizing sialylated oligosaccharides such as ganglioside,

CC lysoganglioside or their mimics. Glycosyltransferases are useful for

CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and

CC other oligosaccharides that have biological activity. The enzymes and

CC nucleic acids that encode them are useful for studies of the pathogenesis

CC mechanisms of organisms that synthesize ganglioside mimics, such as C

CC jejuni and the nucleic acids are used as probes to study expression of

CC genes involved in ganglioside mimetic synthesis. Antibodies raised

CC against the glycosyltransferases are also useful for analyzing the

CC expression patterns of these genes involved in pathogenesis. The nucleic

CC acids are also useful for designing antisense oligonucleotides for

CC inhibiting expression of the Campylobacter enzymes that are involved in

CC the biosynthesis of ganglioside mimics that can mask the pathogens from

CC the host's immune system. The oligosaccharides are useful as diagnosing

CC reagents or as therapeutics and as immunogens for producing antibodies.

CC Bacterial glycosyltransferase can be used to catalyze the formation of





PT New sialyltransferases useful for adding sialyl residues to acceptor  
XX molecules.  
PS Claim 27; Fig 2; 47pp; English.  
XX  
CC The present sequence represents Campylobacter jejuni alpha-2,3-  
CC sialyltransferase which is encoded by the cst-1 gene. The alpha-2,3-  
CC sialyltransferase protein is useful for producing desired carbohydrate  
CC structures by contacting the acceptor molecule (which has a terminal  
CC galactose residue) with an activated sialic acid molecule. The terminal  
CC galactose residue is linked to a second residue (Glc or a GlcNAc, or  
CC GlcNAc or GalNAc) in the acceptor molecule through a beta-1,3 or beta-1,4  
CC linkage, respectively. The activated sialic acid is CMP-Neu5Ac. The  
CC polynucleotides and polypeptides facilitate the improved production of  
CC desired structures and nucleic acids encoding sialyltransferases  
XX  
SQ Sequence 430 AA;

Query Match 47.5%; Score 743; DB 2; Length 430;  
Best Local Similarity 50.5%; Pred. No. 2.6e-65;  
Matches 147; Conservative 45; Mismatches 81; Indels 18; Gaps 5;

QY 1 MKKVIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLGKKCKAVFYTNFFFEQYY 60  
DB 16 MONTIAGNGPSLKNINRYKRPREYDVRCNQFYEDKYLGKKCKAVFFNPGVFLQYH 75

QY 61 TLKHLIQNOYETELIMCSNVAIAALGYKEIYLSGIDFYQSSYAFDTKQENLLKLAP 120  
DB 76 TAKQILKNEVEIKNIFCSTNLPFIESNDFLHQFYNFPPDAKLGVEIENLKEFYAYIK 135

QY 121 PHEIYNQRIITSGVVMCAVAIAALGYKEIYLSGIDFYQSSYAFDTKQENLLKLAP 177  
DB 136 YNEIYFNKRITSGVVMCAIAALGYKTYLGCIDFYEGDVIYFFEAMSTNIKTIFPGIKD 195

QY 178 FKNDRSHVGHKNTDIKALBEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK- 236  
DB 196 FKPSNC-----HSKEYDIEALKLLKLSYKVNITYALCDDSI LANHPFLSINNNFTLENKH 251

QY 237 NNYTKDILIPSEAYGKFSKN-----INFKKI-KIKENVYVKLIKDL 277  
DB 252 NNSINDILLTNTPGVSYFKQKADKNIMLNFNILHSKDWLIKFLNKEI 302

RESULT 14  
ID ABR40195  
XX ABR40195 standard; protein; 303 AA.  
AC ABR40195;  
XX  
XX 23-JUL-2003 (first entry)  
DT alpha-2,3/alpha-2,8-sialyltransferase #3.  
DE  
XX  
XX alpha-2,3/alpha-2,8-sialyltransferase; enzyme; sialic acid.  
KW  
XX  
XX Pasteurella multocida.  
OS  
XX  
XX WO2003027297-A1.  
PN  
XX  
XX 03-APR-2003.  
PD  
XX  
XX 26-SEP-2002; 2002WO-JP009907.  
PF  
XX  
XX 26-SEP-2001; 2001JP-00292796.  
PR  
XX  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX  
XX Endo T, Koizumi S;  
PI  
XX  
XX WPI; 2003-393339/37.  
DR  
XX  
XX N-PSDB; ACC71694.  
DR  
XX  
XX Alpha 2,3/alpha 2,8 sialyltransferase from Pasteurella multocida for

PT production of sialic acid-containing complex sugars.  
XX  
PS Claim 9; Page 60-62; 66pp; Japanese.  
XX  
CC The present invention relates to a method for producing alpha-2,3/alpha-  
CC 2,8-sialyltransferase. alpha-2,3/alpha-2,8-sialyltransferase can be used  
CC in a method for producing sialic acid-containing complex sugars from an  
CC oligosaccharide substrate using a culture of a microorganism transformed  
CC with DNA encoding alpha-2,3/alpha-2,8-sialyltransferase, or an extract of  
CC the culture. The present sequence is a protein sequence for alpha-  
CC 2,3/alpha-2,8-sialyltransferase from Pasteurella multocida  
XX  
SQ Sequence 303 AA;

Query Match 44.3%; Score 692.5; DB 6; Length 303;  
Best Local Similarity 46.1%; Pred. No. 1.8e-60;  
Matches 136; Conservative 58; Mismatches 94; Indels 7; Gaps 5;

QY 2 KKVIIAGNGPSLKEIDYSLRPNDFVRCNQFYEDKYLGKKCKAVFYTNFFFEQYYT 61  
DB 11 KAVIVAGNGESLSQIDYRLLPKNYDVRCNQFYFEERYFLGNKIKAVFFTEGVLEQYYT 70

QY 62 LXHLIQNOYETELIMCSNVAIAALGYKEIYLSGIDFYQSSYAFDTKQENLLKLAP 120  
DB 71 LYHLKENNEYFVDNLTSEFNHPTVDLEKSKQ-IQALFIDVINGEYKLSKLTAPDVYLR 129

QY 121 PHEIYNQRIITSGVVMCAVAIAALGYKEIYLSGIDFYQSSYAFDTKQENLLKLAP 179  
DB 130 YKELIYNQRIITSGVVMCAVAIAALGYTDIYLTGIDFYQASENVAFAFNKNIIRLLPDR 189

QY 180 NDRSHVGHKNTDIKALBEFLEKTYKIKLYCLCPNSLLANFIEL---APNLSNFIIOEK 236  
DB 190 KETLPSYHSKDIIDLEALGFLQOHYHWFYSISPMSPSLSKHFPPTVEDDCETTFVAPLK 249

QY 237 NNYTKDILIPSEAYGKFSKNINFKIKIKENVYVKLIKDLRLPSDIKHFKGK 291  
DB 250 ENYINDILLPPHFVYEKLGITVS-KKSRFHSNLTIVRLIRDLKLPSALKHYLKEK 303

RESULT 15  
ID ADN19559  
XX ADN19559 standard; protein; 958 AA.  
AC ADN19559;  
XX  
XX 02-DEC-2004 (first entry)  
DT  
XX  
XX Bacterial polypeptide #2212.  
DE  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
XX Bacteria.  
OS  
XX  
XX US2003233675-A1.  
PN  
XX  
XX 18-DEC-2003.  
PD  
XX  
XX 20-FEB-2003; 2003US-00369493.  
PF  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
PR  
XX  
XX (CAOY/) CAO Y.  
PA (HINK/) HINKLE G J.  
PA (SLAT/) SLATER S C.  
PA (CHEN/) CHEN X.  
PA (GOLD/) GOLDMAN B S.  
XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 2212; 122bp; English.

XX The invention relates to a recombinant DNA construct comprising a  
CC promoter functional in a plant cell, where the promoter is positioned to  
CC provide for expression of a polynucleotide encoding a polypeptide from a  
CC microbial source. The invention also relates to a transformed plant  
CC comprising the recombinant DNA construct and a method of producing a  
CC transformed plant having an improved property. The plant is a crop plant  
CC such as maize or soybean. The method of producing a transformed plant  
CC having an improved property comprises transforming a plant with the  
CC recombinant DNA construct and growing the transformed plant, where the  
CC polynucleotide or polypeptide is useful for improving plant properties.  
CC The recombinant DNA construct is useful for producing plants with  
CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
CC increased resistance to plant disease, better growth rate by modification  
CC of the cell cycle pathway with plant growth regulators, increased rate of  
CC homologous recombination, modified seed oil or protein yield and/or  
CC content, improved yield by modification of carbohydrate, nitrogen or  
CC phosphorus use and/or uptake, by modification of photosynthesis or by  
CC providing improved plant growth and development under at least one stress  
CC condition, improved lignin production or improved galactomannan  
CC production. This sequence represents a bacterial polypeptide used in the  
CC scope of the invention. Note: The sequence data for this patent did not  
CC form part of the printed specification but was obtained in electronic  
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 958 AA;

Query Match 7.6%; Score 119.5; DB 8; Length 958;  
Best Local Similarity 23.0%; Pred. No. 0.025;  
Matches 70; Conservative 40; Mismatches 96; Indels 99; Gaps 18;  
QY 41 LGK----KCAVFTPNFFPEQYTYLKLH-----IQOEYETELIMCSYNOAHLENF 91  
DB 578 LGKGNFGKWLAEYKVN---KKFYAIKVLKKEAILKNEELESKTEKHVFVANKEHFF 634  
QY 92 VKTFDYDYPDAH-----LGVDFFKQKE-----FNAYKFEHLYFNQRI 131  
DB 635 LLNLFASFQSTRVYFVMEYILGGDLMVHIOQQQFSVKRARXYFNLYFYF---YFYPLLT 691  
QY 132 SGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLIKAPDFKNDRSHVIGHSKN 191  
DB 692 PXFYGAIEVCLAKY-----PHENGIAVR-DIKLDNL-LCPD-----GHIRI 731  
QY 192 TDIKALFLEKTYKIKYLC-C-PNSLLANFIE-----LAPNLNSN---FIIQEK-- 237  
DB 732 AD-----YGLCKENMLLGNLTSTFCGTPEFMAPEVRRLFWFILSNENRF 776  
QY 238 ----NYTKDILIPSEAYGKFNKFKIK-----ENVYKLIKD-----LLRLPSDIK 285  
DB 777 LLEQQYSKDV---DWWAFGLVMYQMLIGQSPFGKEDEEIPDAILSDPELPFINNPADAV 833  
QY 286 HYFKG 290  
DB 834 SLIRG 838

Search completed: July 19, 2006, 13:37:33  
Job time : 108.972 secs

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2006, 21:54:06 ; Search time 8431 Seconds  
(without alignments)  
6644.282 Million cell updates/sec

Title: US-10-734-719-8  
Perfect score: 876  
Sequence: 1 atgaaaaaagtattattgc.....attattcaaggaaataaa 876

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 500 summaries

Database : GenEmbl.\*

- 1: gb\_env.\*
- 2: gb\_pat.\*
- 3: gb\_ph.\*
- 4: gb\_pl.\*
- 5: gb\_pr.\*
- 6: gb\_ro.\*
- 7: gb\_sta.\*
- 8: gb\_sy.\*
- 9: gb\_un.\*
- 10: gb\_vi.\*
- 11: gb\_ov.\*
- 12: gb\_hvg.\*
- 13: gb\_in.\*
- 14: gb\_om.\*
- 15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	876	100.0	876	2	AX934431 Sequence
2	876	100.0	876	2	BD249794 Campyloba
3	876	100.0	876	2	AR271703 Sequence
4	876	100.0	876	2	AR481785 Sequence
5	876	100.0	876	2	AR527384 Sequence
6	876	100.0	876	2	AR609663 Sequence
7	876	100.0	876	2	AR689941 Sequence
8	876	100.0	876	2	AR691837 Sequence
9	876	100.0	6047	15	AY661458 Campyloba
10	876	100.0	24425	15	AY422197 Campyloba
11	876	100.0	24437	15	AF167344 Campyloba
12	874.4	99.8	4749	15	AF305571 Campyloba
13	872.8	99.6	12390	15	AF401528 Campyloba
14	869.8	99.3	873	2	AX934436 Sequence
15	869.8	99.3	873	2	AR481787 Sequence
16	869.8	99.3	873	2	AR527386 Sequence
17	869.8	99.3	873	2	AR609665 Sequence
18	869.8	99.3	873	2	AR689943 Sequence

19	869.8	99.3	873	2	AR691839 Sequence
20	869.6	99.3	11474	15	AF215659 Campyloba
21	866.6	98.9	873	2	AX934434 Sequence
22	866.6	98.9	873	2	AR481786 Sequence
23	866.6	98.9	873	2	AR527385 Sequence
24	866.6	98.9	873	2	AR609664 Sequence
25	866.6	98.9	873	2	AR689942 Sequence
26	866.6	98.9	873	2	AR691838 Sequence
27	852.4	97.3	1303	15	AY681242 Campyloba
28	844	96.3	876	2	AX934427 Sequence
29	844	96.3	876	2	BD249792 Campyloba
30	844	96.3	876	2	AR271701 Sequence
31	844	96.3	876	2	AR481783 Sequence
32	844	96.3	876	2	AR527382 Sequence
33	844	96.3	876	2	AR609661 Sequence
34	844	96.3	876	2	AR689939 Sequence
35	844	96.3	876	2	AR691835 Sequence
36	842.4	96.2	11442	15	AF400048 Campyloba
37	837.6	95.6	876	15	AF216647 Campyloba
38	837.6	95.6	6047	15	AY644679 Campyloba
39	837.6	95.6	11474	2	AX934424 Sequence
40	837.6	95.6	11474	2	BD249790 Campyloba
41	837.6	95.6	11474	2	AR271699 Sequence
42	837.6	95.6	11474	2	AR481781 Sequence
43	837.6	95.6	11474	2	AR527380 Sequence
44	837.6	95.6	11474	2	AR609659 Sequence
45	837.6	95.6	11474	2	AR689937 Sequence
46	837.6	95.6	11474	2	AR691833 Sequence
47	837.6	95.6	11474	15	AF130984 Campyloba
48	837.6	95.6	12576	15	AY297047 Campyloba
49	836	95.4	876	2	AX934425 Sequence
50	836	95.4	876	2	BD249791 Campyloba
51	836	95.4	876	2	AR271700 Sequence
52	836	95.4	876	2	AR481782 Sequence
53	836	95.4	876	2	AR527381 Sequence
54	836	95.4	876	2	AR609660 Sequence
55	836	95.4	876	2	AR689938 Sequence
56	836	95.4	876	2	AR691834 Sequence
57	820	93.6	876	2	AX934429 Sequence
58	820	93.6	876	2	BD249793 Campyloba
59	820	93.6	876	2	AR271702 Sequence
60	820	93.6	876	2	AR481784 Sequence
61	820	93.6	876	2	AR527383 Sequence
62	820	93.6	876	2	AR609662 Sequence
63	820	93.6	876	2	AR689940 Sequence
64	820	93.6	876	2	AR691836 Sequence
65	818.4	93.4	11455	15	AY044868 Campyloba
66	805.6	92.0	12388	15	AF401529 Campyloba
67	780.8	89.1	12403	15	AY854153 Campyloba
68	779.8	89.0	12370	15	AY422196 Campyloba
69	311.6	35.6	1293	2	BD134499 Campyloba
70	311.6	35.6	1293	2	AR473907 Sequence
71	311.6	35.6	1293	2	AR489130 Sequence
72	311.6	35.6	10484	15	AY791515 Campyloba
73	311.6	35.6	11164	15	AF130466 Campyloba
74	306	34.9	3071	15	AF195055 Campyloba
75	306	34.9	4668	15	AF257460 Campyloba
76	306	34.9	13484	15	AF400047 Campyloba
77	306	34.9	13484	15	AY044156 Campyloba
78	306	34.9	282183	15	CJ11168X4 Campyloba
79	286.6	32.7	110000	2	BD426631_03 Continuation (4 of
80	286.6	32.7	110000	2	AR274511_03 Continuation (4 of
81	286.6	32.7	110000	2	AR632719_03 Continuation (4 of
82	286.6	32.7	110000	15	L42023_03 Continuation (10 o
83	281.8	32.2	110000	15	CP000057_09 CQ873070 Sequence
84	281.8	32.2	191996	2	CQ873070 Sequence
85	265.4	30.3	960	2	CS222474 Sequence
86	265.4	30.3	13379	2	CQ872953 Sequence
87	265.4	30.3	110000	15	CP000057_04 Continuation (5 of
88	265.4	30.3	110000	15	CQ873159 Sequence
89	265.4	30.3	349980	2	CQ873071 Sequence
90	260.4	29.7	4277	2	CQ872844 Sequence
91	257.8	29.4	903	2	CS220684 Sequence

C 92	257	29.3	10024	15	AE006157	AE006157 Pasteurel	165	77.6	8.9	110000	4	AP007151_00	AP007151 Aspergill
C 93	178.2	20.3	4977	15	HILC3	X57315 Haemophilus	C 166	77.4	8.8	2009	2	AX457067_00	AX457067 Sequence
C 94	99.4	11.3	60604	12	AC023466	AC023466 Homo sapi	C 167	77.4	8.8	130012	12	BX957280	BX957280 Danio rer
C 95	95.8	10.9	34980	12	AX344555	AX344555 Sequence	C 168	77.4	8.8	185857	12	CR847995	CR847995 Danio rer
C 96	92.2	10.5	231489	12	AC156865	AC156865 Bos tauru	C 169	77.2	8.8	152209	5	HS1108D11	AL034419 Human DNA
C 97	91.2	10.4	263707	12	CR356223	CR356223 Danio rer	C 170	77.2	8.8	163735	5	AC027277	AC027277 Homo sapi
C 98	90.2	10.3	2482	2	CS084032	CS084032 Sequence	C 171	77.2	8.8	194028	12	AC111077	AC111077 Homo sapi
C 99	90.2	10.3	122168	2	AC127383	AC127383 Homo sapi	C 172	77.2	8.8	196103	5	BX842591	BX842591 Rhesus DN
C 100	89	10.2	1192	5	HA323759	HA323759 Homo sapi	C 173	77	8.8	110000	12	AC108692_0	AC108692 Homo sapi
C 101	89	10.2	8056	2	AX599046	AX599046 Sequence	C 174	77	8.8	186431	12	AC022281	AC022281 Homo sapi
C 102	88.8	10.1	8056	2	AX599046	AX599046 Sequence	C 175	76.8	8.8	154071	13	AC115598	AC115598 Dictyoste
C 103	88.8	10.1	180629	12	CR388025	CR388025 Danio rer	C 176	76.8	8.8	162245	11	BX571771	BX571771 Zebrafish
C 104	88.8	10.1	185596	5	AC021553	AC021553 Homo sapi	C 177	76.8	8.8	254436	13	AE014827	AE014827 Plasmodiu
C 105	87.6	10.0	196784	5	AC016543	AC016543 Homo sapi	C 178	76.8	8.8	347582	13	PFMAL4P1	AL034557 Plasmodiu
C 106	86.2	9.8	19087	2	AX345695	AX345695 Sequence	C 179	76.8	8.8	349980	2	AX344560	AX344560 Plasmodiu
C 107	86.2	9.8	19087	2	AX345695	AX345695 Sequence	C 180	76.6	8.7	6465	2	AX345887	AX345887 Sequence
C 108	85.4	9.7	202971	12	CR759948	CR759948 Danio rer	C 181	76.6	8.7	104992	12	AC005504	AC005504 Plasmodiu
C 109	84.8	9.7	148833	12	CR753884	CR753884 Danio rer	C 182	76.6	8.7	110000	12	PFMAL7P1_07	Continuation (8 of
C 110	84.6	9.7	110000	12	PFMAL8P1_10	Continuation (11 o	C 183	76.6	8.7	110350	12	CT573346	CT573346 Danio rer
C 111	84.6	9.7	348174	13	CR382399	CR382399 Plasmodiu	C 184	76.6	8.7	161757	12	AC044901	AC044901 Homo sapi
C 112	84.4	9.6	86826	13	PFMAL3P5	AL034556 Plasmodiu	C 185	76.6	8.7	169546	12	AC004157	AC004157 Plasmodiu
C 113	84.2	9.6	86826	13	PFMAL3P5	AL034556 Plasmodiu	C 186	76.6	8.7	250421	13	AE014849	AE014849 Plasmodiu
C 114	84	9.6	221958	12	CR847903	CR847903 Danio rer	C 187	76.4	8.7	76568	13	MBREV	AF538053 Monosiga
C 115	83.6	9.5	154563	5	CR936360	CR936360 Human DNA	C 188	76.4	8.7	90373	13	AC115680	AC115680 Dictyoste
C 116	83.2	9.5	250029	13	AE014839	AE014839 Plasmodiu	C 189	76.4	8.7	110000	13	AC116984_2	Continuation (3 of
C 117	83	9.5	110000	12	PFMAL13_23	Continuation (24 o	C 190	76.4	8.7	135370	12	CT027572	CT027572 Danio rer
C 118	83	9.5	348174	13	CR382399	CR382399 Plasmodiu	C 191	76.4	8.7	153499	11	CR790375	CR790375 Zebrafish
C 119	82.8	9.5	114276	5	AC011355	AC011355 Homo sapi	C 192	76.4	8.7	242222	5	AC099668	AC099668 Homo sapi
C 120	82.8	9.5	175544	12	AC117342	AC117342 Rattus no	C 193	76.2	8.7	164241	11	CR383676	CR383676 Zebrafish
C 121	82.6	9.4	8056	2	AX598900	AX598900 Sequence	C 194	76	8.7	17131	2	AX345955	AX345955 Sequence
C 122	82.4	9.4	250237	12	AC160076	AC160076 Bos tauru	C 195	76	8.7	135121	5	AC069525	AC069525 Homo sapi
C 123	82	9.4	144493	5	AP001547	AP001547 Homo sapi	C 196	76	8.7	225584	12	AC168241	AC168241 Bos tauru
C 124	82	9.4	349751	13	PFMAL4P3	AL035476 Plasmodiu	C 197	76	8.7	335050	13	PFPA923356	AL929356 Plasmodiu
C 125	81.8	9.3	176898	5	AC117569	AC117569 Homo sapi	C 198	75.8	8.7	110000	5	AL731858	AL731858 Human DNA
C 126	81.6	9.3	110000	4	AP007151_00	AP007151 Aspergill	C 199	75.8	8.7	182518	5	AC131649	AC131649 Homo sapi
C 127	81.4	9.3	30157	5	AC073242	AC073242 Homo sapi	C 200	75.8	8.7	214035	12	CR848786	CR848786 Danio rer
C 128	81.4	9.3	162445	5	AL158151	AL158151 Human DNA	C 201	75.8	8.7	253001	13	AE014834	AE014834 Plasmodiu
C 129	81.2	9.3	131682	5	AL672277	AL672277 Human DNA	C 202	75.8	8.7	253001	13	AE014834	AE014834 Plasmodiu
C 130	81.2	9.3	200187	12	CR925711	CR925711 Danio rer	C 203	75.6	8.6	1501	2	AX598988	AX598988 Sequence
C 131	81	9.2	180861	12	CR391970	CR391970 Danio rer	C 204	75.6	8.6	1501	2	AX767510	AX767510 Sequence
C 132	81	9.2	313050	13	PFPA929352	AL929352 Plasmodiu	C 205	75.6	8.6	141469	5	AL139811	AL139811 Human DNA
C 133	80.8	9.2	86479	11	CR936500	CR936500 Zebrafish	C 206	75.6	8.6	163443	12	AC006280	AC006280 Plasmodiu
C 134	80.6	9.2	147340	12	CR854842	CR854842 Danio rer	C 207	75.6	8.6	225584	12	AC168241	AC168241 Bos tauru
C 135	80.4	9.2	15548	2	AX347057	AX347057 Sequence	C 208	75.6	8.6	252850	13	AE014847	AE014847 Plasmodiu
C 136	80.2	9.2	110000	12	TANN3_01	Continuation (2 of	C 209	75.4	8.6	104992	12	AC005504	AC005504 Plasmodiu
C 137	80	9.1	212999	12	AC151201	AC151201 Bos tauru	C 210	75.4	8.6	130141	11	BX640583	BX640583 Zebrafish
C 138	79.8	9.1	175946	5	AC141000	AC141000 Homo sapi	C 211	75.4	8.6	155648	12	CT573493	CT573493 Danio rer
C 139	79.6	9.1	1554	2	CS083940	CS083940 Sequence	C 212	75.4	8.6	169546	12	AC004157	AC004157 Plasmodiu
C 140	79.6	9.1	110000	4	AP007164_00	AP007164 Aspergill	C 213	75.4	8.6	178785	5	AC068139	AC068139 Homo sapi
C 141	79.6	9.1	170627	12	AC125567	AC125567 Rattus no	C 214	75.4	8.6	182155	5	AC104814	AC104814 Homo sapi
C 142	79.6	9.1	183584	5	AC012492	AC012492 Homo sapi	C 215	75.4	8.6	184838	5	AC090740	AC090740 Homo sapi
C 143	79.6	9.1	252632	13	AE014818	AE014818 Plasmodiu	C 216	75.4	8.6	250421	13	AE014849	AE014849 Plasmodiu
C 144	79.4	9.1	67970	13	PFMAL1P3	AL031746 Plasmodiu	C 217	75.2	8.6	83435	11	BX640469	BX640469 Zebrafish
C 145	79.2	9.0	3206	13	AB242844	AB242844 Drosophil	C 218	75.2	8.6	150199	5	AC087568	AC087568 Pan trogl
C 146	79.2	9.0	110000	12	PFMAL13_16	Continuation (17 o	C 219	75.2	8.6	245106	12	CEY53C12D	AL033123 Caenorhab
C 147	79.2	9.0	110000	12	PFMAL13_17	Continuation (18 o	C 220	75	8.6	1496	13	AY392440	AY392440 Dictyoste
C 148	79.2	9.0	229210	12	CR847561	CR847561 Danio rer	C 221	75	8.6	110000	12	PFMAL8P1_04	Continuation (5 of
C 149	79	9.0	60282	12	CR933812	CR933812 Danio rer	C 222	75	8.6	163964	11	BX649330	BX649330 Zebrafish
C 150	79	9.0	86826	13	PFMAL3P5	AL034556 Plasmodiu	C 223	75	8.6	226857	12	CT573293	CT573293 Danio rer
C 151	78.8	9.0	4601	13	DMU11584	U11584 Drosophila	C 224	75	8.6	26555	12	CT573293	CT573293 Danio rer
C 152	78.8	9.0	19517	13	DMU37541	U37541 Drosophila	C 225	74.8	8.5	78578	12	PFMAL13P9	AL096783 Plasmodiu
C 153	78.8	9.0	28717	5	AL513226	AL513226 Human DNA	C 226	74.8	8.5	96799	5	AC104647	AL04647 Homo sapi
C 154	78.8	9.0	108902	12	AC011430	AC011430 Homo sapi	C 227	74.8	8.5	96799	5	AC117444	AC117444 Homo sapi
C 155	78.6	9.0	91798	5	HSXB152G3	AL132982 Human DNA	C 228	74.8	8.5	104853	5	AP007164	AP007164 Aspergill
C 156	78.6	9.0	267002	12	CR354431	CR354431 Danio rer	C 229	74.8	8.5	110000	4	AP007164_00	Continuation (5 of
C 157	78.2	8.9	99263	12	CNS01DX9	AL139177 Homo sapi	C 230	74.8	8.5	110000	12	CT009752_4	Continuation (6 of
C 158	78.2	8.9	105940	5	AC010606	AC010606 Homo sapi	C 231	74.8	8.5	110000	12	CT009752_5	Continuation (16 o
C 159	78	8.9	8056	2	AX598900	AX598900 Sequence	C 232	74.8	8.5	110000	12	PFMAL13_15	BX248123 Human DNA
C 160	78	8.9	110000	12	AC108692_0	AC108692 Homo sapi	C 233	74.8	8.5	134627	5	BX248123	AC027714 Homo sapi
C 161	78	8.9	145889	12	CT573340	CT573340 Danio rer	C 234	74.8	8.5	154289	12	AC027714	AC027714 Homo sapi
C 162	78	8.9	186431	12	AC022281	AC022281 Homo sapi	C 235	74.8	8.5	184039	5	AC099327	AC099327 Homo sapi
C 163	77.8	8.9	14867	13	AE001398	AE001398 Plasmodiu	C 236	74.8	8.5	184704	12	CT025866	CT025866 Danio rer
C 164	77.8	8.9	88037	12	PFMAL13P8	AL096782 Plasmodiu	C 237	74.6	8.5	4120	13	DQ054839	DQ054839 Plasmodiu

c 238	74.6	8.5	94534	11	AL929250	AL929250 Zebrafish	311	73	8.3	106763	5	AP002091	AP002091 Homo sapi
239	74.6	8.5	111882	13	AC115612	Dictyoste	312	73	8.3	164686	5	CNS05TCG	AL355097 Human chr
240	74.6	8.5	156794	12	CT027830	Danio rer	313	73	8.3	178207	5	AL140172	AL140172 Homo sapi
c 241	74.6	8.5	180914	12	CT033820	Danio rer	c 314	73	8.3	179603	2	CO870485	CO870485 Sequence
242	74.6	8.5	182195	12	CT573300	Danio rer	c 315	73	8.3	249943	13	AE014823	AE014823 Plasmid
c 243	74.6	8.5	206709	12	CR762441	Danio rer	316	73	8.3	254050	13	PF929358	AL929358 Plasmid
c 244	74.4	8.5	1090	13	AB084761	Drosophila	317	73	8.3	268322	12	AC169256	AL169256 Bos tauru
c 245	74.4	8.5	76568	13	MBREV	AF538053 Monogiga	c 318	72.8	8.3	15421	13	PF00MPIRA	X95275 Plasmid
c 246	74.4	8.5	176758	12	CR352258	Danio rer	c 319	72.8	8.3	175559	12	AC145101	AX344573 Sequence
c 247	74.2	8.5	133877	12	AC120883	Danio rer	c 320	72.8	8.3	349980	2	AX344573	AX344573 Sequence
248	74.2	8.5	137889	5	AC073269	AC073269 Homo sapi	c 321	72.6	8.3	5689	2	AX251180	AX251180 Sequence
249	74.2	8.5	141721	12	AC149953	AC149953 Strongylo	c 322	72.6	8.3	5689	2	AX277927	AX277927 Sequence
250	74.2	8.5	167326	11	CR354435	Zebrafish	c 323	72.6	8.3	5689	2	AX323612	AX323612 Sequence
c 251	74.2	8.5	169510	12	CR855864	Danio rer	324	72.6	8.3	15421	13	PF00MPIRA	X95275 Plasmid
c 252	74.2	8.5	175559	12	AC145101	Danio rer	c 325	72.6	8.3	175559	12	AC145101	AX344573 Sequence
c 253	74	8.4	18988	2	AX251096	AX251096 Sequence	326	72.6	8.3	349980	2	AX344573	AX344573 Sequence
c 254	74	8.4	18988	2	AX281398	AX281398 Sequence	327	72.6	8.3	5689	2	AX251180	AX251180 Sequence
c 255	74	8.4	18988	2	AX345603	AX345603 Sequence	c 328	72.6	8.3	130851	11	AX664721	AX664721 Zebrafish
c 256	74	8.4	18988	2	AX348636	AX348636 Sequence	329	72.6	8.3	134708	5	AC147046	AC147046 Pan trogl
c 257	74	8.4	99003	12	AL3390756	Continuation (5 of	330	72.6	8.3	145580	5	AC097373	AC097373 Homo sapi
c 258	74	8.4	104369	4	AF007170_4	Continuation (5 of	331	72.6	8.3	203204	5	AC092052	AC092052 Homo sapi
c 259	74	8.4	110000	12	PFMAL13_20	Continuation (21 o	332	72.6	8.3	243899	11	EX510919	EX510919 Zebrafish
c 260	74	8.4	110000	13	AC116957_0	AC116957 Dictyoste	333	72.6	8.3	252125	12	CR848052	CR848052 Danio rer
261	74	8.4	116344	12	CR962131_1	CR962131 Medicago	c 334	72.6	8.3	254050	13	PF929358	AL929358 Plasmid
262	74	8.4	171403	5	AC092701	AC092701 Homo sapi	c 335	72.6	8.3	349210	12	CR396591	CR396591 Danio rer
263	74	8.4	190670	11	EX000521	EX000521 Zebrafish	c 336	72.6	8.3	349751	13	PFMAL4P3	AL035476 Plasmid
c 264	74	8.4	338231	4	AF007163	AP007163 Aspergill	c 337	72.4	8.3	3241	13	AX691413	AX691413-Erium glo
c 265	73.8	8.4	5930	2	AX345419	AX345419 Sequence	c 338	72.4	8.3	5979	2	AX277855	AX277855 Sequence
c 266	73.8	8.4	14867	13	AE001398	AE001398 Plasmid	c 339	72.4	8.3	5979	2	AX277855	AX277855 Sequence
c 267	73.8	8.4	22368	13	AY293825	AY293825 Dictyoste	c 340	72.4	8.3	6132	2	AX345765	AX345765 Sequence
c 268	73.8	8.4	105238	5	AC011458	AC011458 Homo sapi	341	72.4	8.3	87741	11	CR788248	CR788248 Zebrafish
c 269	73.8	8.4	115990	11	EX890614	EX890614 Zebrafish	c 342	72.4	8.3	99089	4	AP007265	AP007265 Lotus cor
c 270	73.8	8.4	143331	5	AC091214	AC091214 Homo sapi	343	72.4	8.3	110000	13	PFMAL1P2_2	Continuation (3 of
271	73.8	8.4	174631	12	AC069059	AC069059 Homo sapi	344	72.4	8.3	110350	12	CT573346	CT573346 Danio rer
c 272	73.8	8.4	221775	12	CT027701	CT027701 Danio rer	c 345	72.4	8.3	163149	5	CNS01DSZ	AL122057 Human chr
c 273	73.8	8.4	253305	13	PFMAL3P7	AL034559 Plasmid	c 346	72.4	8.3	177421	5	CR295781	CR295781 Danio rer
c 274	73.8	8.4	256879	13	AC116982	AC116982 Dictyoste	c 347	72.4	8.3	204819	5	AC024934	AC024934 Homo sapi
c 275	73.8	8.4	302156	13	AC116977	AC116977 Dictyoste	348	72.4	8.3	250029	13	AE014839	AE014839 Plasmid
c 276	73.6	8.4	961	13	AX351412	AX351412 Aethusan	349	72.4	8.3	250195	13	AE014831	AE014831 Plasmid
c 277	73.6	8.4	1824	15	AF267217	AF267217 Candidatu	350	72.4	8.3	348034	13	CR382400	CR382400 Plasmid
c 278	73.6	8.4	53932	12	AC023371	AC023371 Homo sapi	c 351	72.2	8.2	34996	13	U87145	U87145 Toxoplasma
c 279	73.6	8.4	10150	5	AC134620	AC134620 Homo sapi	c 352	72.2	8.2	110000	15	BA000021_3	Continuation (4 of
280	73.6	8.4	108865	5	HS0345E4	AL132661 Human DNA	c 353	72.2	8.2	136698	12	CT573297	CT573297 Danio rer
281	73.6	8.4	110000	12	PFMAL8P1_04	Continuation (5 of	c 354	72.2	8.2	142372	5	AC132152	AC132152 Homo sapi
282	73.6	8.4	110000	12	PFMAL8P1_05	Continuation (6 of	c 355	72.2	8.2	178193	12	AC068949	AC068949 Homo sapi
c 283	73.6	8.4	187264	12	CR847945	CR847945 Danio rer	c 356	72.2	8.2	192265	5	CNS018P3	AL10118 Human chr
c 284	73.6	8.4	192731	5	AC026698	AC026698 Homo sapi	c 357	72.2	8.2	224497	12	CR954145	CR954145 Danio rer
c 285	73.6	8.4	205429	12	AC005506	AC005506 Plasmid	c 358	72.2	8.2	935	13	AX351413	AX351413 Aethusan
c 286	73.6	8.4	253132	13	AE014846	AE014846 Plasmid	c 359	72.2	8.2	969	13	DMA132902	AD132902 Drosophila
c 287	73.6	8.4	330767	12	AC167014	AC167014 Bos tauru	c 360	72.2	8.2	89357	5	AC109472	AC109472 Homo sapi
288	73.6	8.4	347050	13	PF929351	AL929351 Plasmid	c 361	72.2	8.2	175664	5	AC027811	AC027811 Homo sapi
c 289	73.4	8.4	12592	2	AX251797	AX251797 Sequence	c 362	72.2	8.2	181792	5	AC098822	AC098822 Homo sapi
c 290	73.4	8.4	66934	11	BX539335	BX539335 Zebrafish	c 363	72.2	8.2	210916	12	CT573319	CT573319 Danio rer
c 291	73.4	8.4	110000	4	CR382132_17	Continuation (18 o	c 364	72.2	8.2	251448	13	AE014819	AE014819 Plasmid
c 292	73.4	8.4	145992	5	AC134919	AC134919 Homo sapi	365	71.8	8.2	1192	5	HS4323759	AT323759 Homo sapi
c 293	73.4	8.4	150101	11	EX248231	EX248231 Zebrafish	366	71.8	8.2	44352	13	AF098501	AF098501 Caenorhab
294	73.4	8.4	157210	11	EX248231	EX248231 Zebrafish	367	71.8	8.2	135121	5	AC069525	AC069525 Homo sapi
c 295	73.4	8.4	177650	5	AC104027	AC104027 Homo sapi	c 368	71.8	8.2	156295	11	BX000703	BX000703 Zebrafish
c 296	73.4	8.4	182518	5	AC131649	AC131649 Homo sapi	369	71.8	8.2	166719	5	AC136957	AC136957 Homo sapi
297	73.4	8.4	209157	5	CNS01DW4	AL136332 Human chr	c 370	71.8	8.2	182195	12	CT573300	CT573300 Danio rer
298	73.4	8.4	247919	12	AC151416	AC151416 Bos tauru	c 371	71.8	8.2	185640	11	AL772314	AL772314 Zebrafish
c 299	73.2	8.4	36977	12	AC092304	AC092304 Homo sapi	c 372	71.8	8.2	186875	5	AL731540	AL731540 Human DNA
300	73.2	8.4	84563	11	AC096885	AC096885 Danio rer	c 373	71.8	8.2	196988	5	AP002761	AP002761 Homo sapi
301	73.2	8.4	126999	5	AL513328	AL513328 Human DNA	374	71.8	8.2	198434	5	AC109994	AC109994 Homo sapi
302	73.2	8.4	144604	12	EX899179	EX899179 Danio rer	c 375	71.8	8.2	257109	13	AC116963	AC116963 Dictyoste
303	73.2	8.4	145553	11	AL935272	AL935272 Zebrafish	c 376	71.6	8.2	18166	12	PSMAL8PE	AL929364 Plasmid
304	73.2	8.4	158947	12	AC012220	AC012220 Homo sapi	c 377	71.6	8.2	106763	5	AP002091	AP002091 Homo sapi
c 305	73.2	8.4	162831	5	AC106724	AC106724 Homo sapi	378	71.6	8.2	142372	5	AC132152	AC132152 Homo sapi
c 306	73.2	8.4	164598	5	AC117384	AC117384 Homo sapi	c 379	71.6	8.2	147114	5	BS000121	BS000121 Pan trogl
307	73.2	8.4	201659	12	AC146419	AC146419 Pan trogl	c 380	71.6	8.2	206784	4	HS93L7	AL022401 Human DNA
308	73.2	8.4	250663	13	AE014826	AE014826 Plasmid	381	71.6	8.2	225236	12	BX936371	AX936371 Danio rer
309	73.2	8.4	258658	13	AE014832	AE014832 Plasmid	c 382	71.6	8.2	226345	5	AC005406	AC005406 Homo sapi
310	73	8.3	29080	4	AP007151_44	Continuation (45 o	c 383	71.6	8.2	250078	13	AE014829	AE014829 Plasmid

384	71.6	8.2	250713	13	AE014850	AE014850 Plasmodiu	C 457	70.2	8.0	1791	13	AF151389	AF151389 Dermatobi
C 385	71.6	8.2	349402	15	BX842644	BX842644 Mycoplasma	C 458	70.2	8.0	2215	13	S39048	S39048 knob associ
C 386	71.4	8.2	193800	2	AX252129	AX252129 Sequence	C 459	70.2	8.0	4471	2	CS124698	CS124698 Sequence
C 387	71.4	8.2	94595	5	AL732591	AL732591 Human DNA	C 460	70.2	8.0	8201	2	AX281285	AX281285 Sequence
388	71.4	8.2	110000	12	PFMAL13_05	Continuation (6 of	C 461	70.2	8.0	8201	2	AX345208	AX345208 Sequence
389	71.4	8.2	110000	12	TANNA_01	Continuation (2 of	C 462	70.2	8.0	13286	2	CQ807296	CQ807296 Sequence
C 390	71.4	8.2	162445	11	AL158151	AL158151 Human DNA	C 463	70.2	8.0	25753	4	AF437291	AF437291 Saccharom
C 391	71.4	8.2	167918	5	BX936450	BX936450 Zebrafish	C 464	70.2	8.0	37895	12	AC146472	AC146472 Hylobates
C 392	71.4	8.2	168212	11	CR352223	CR352223 Zebrafish	C 465	70.2	8.0	46930	4	AX534144	AX534144 Saprolign
C 393	71.4	8.2	175842	5	AL731547	AL731547 Human DNA	C 466	70.2	8.0	110000	12	BX322593_4	Continuation (5 of
C 394	71.4	8.2	17893	5	AC005089	AC005089 Homo sapi	C 467	70.2	8.0	118642	5	AC126283	AC126283 Homo sapi
C 395	71.4	8.2	192929	12	AC005505	AC005505 Plasmodiu	C 468	70.2	8.0	130110	5	AL731777	AL731777 Human DNA
C 396	71.4	8.2	195369	11	BX663499	BX663499 Zebrafish	C 469	70.2	8.0	158537	11	BX284640	BX284640 Zebrafish
C 397	71.4	8.2	203320	12	AC136933	AC136933 Homo sapi	C 470	70.2	8.0	158537	11	CT573343	CT573343 Danio rer
C 398	71.4	8.2	337203	13	CR382401	CR382401 Plasmodiu	C 471	70.2	8.0	167701	12	CT573343	CT573343 Danio rer
C 399	71.2	8.1	15723	13	AY521251	AY521251 Aleurodic	C 472	70.2	8.0	177893	5	AC005089	AC005089 Homo sapi
C 400	71.2	8.1	17738	2	AX346441	AX346441 Sequence	C 473	70	8.0	236418	12	CR318626	CR318626 Danio rer
C 401	71.2	8.1	19051	5	AC144547	AC144547 Homo sapi	C 474	70	8.0	3254	13	AY691419	AY691419 Planococc
C 402	71.2	8.1	113201	5	AC117509	AC117509 Homo sapi	C 475	70	8.0	6314	2	AX281272	AX281272 Sequence
C 403	71.2	8.1	116344	12	CR962131	CR962131 Medicago	C 476	70	8.0	16665	13	AY072044	AY072044 Aedes alb
C 404	71.2	8.1	118642	5	AC126283	AC126283 Homo sapi	C 477	70	8.0	84519	12	AC180269	AC180269 Strongylo
C 405	71.2	8.1	132943	5	AC011287	AC011287 Homo sapi	C 478	70	8.0	110000	12	TANN2_02	Continuation (3 of
C 406	71.2	8.1	139917	11	BX510640	BX510640 Zebrafish	C 479	70	8.0	14724	12	CR759566	CR759566 Danio rer
C 407	71.2	8.1	154071	13	AC115598	AC115598 Dictyoste	C 480	70	8.0	148435	5	AC073973	AC073973 Homo sapi
C 408	71.2	8.1	157062	5	AL663105	AL663105 Human DNA	C 481	70	8.0	155885	12	CR848672	CR848672 Danio rer
C 409	71.2	8.1	184542	12	AC150448	AC150448 Pan trogl	C 482	70	8.0	170781	5	AC013473	AC013473 Homo sapi
C 410	71.2	8.1	195971	12	CT027642	CT027642 Danio rer	C 483	70	8.0	180661	5	BS000013	BS000013 Pan trogl
C 411	71.2	8.1	202366	12	BX927344	BX927344 Danio rer	C 484	70	8.0	192372	12	AC110301	AC110301 Homo sapi
C 412	71.2	8.1	253924	13	AE014842	AE014842 Plasmodiu	C 485	70	8.0	196472	12	CT025893	CT025893 Danio rer
C 413	71.2	8.1	271546	13	AE014843	AE014843 Plasmodiu	C 486	70	8.0	201535	5	AC007248	AC007248 Homo sapi
C 414	71.2	8.1	341050	13	PF929357	PF929357 Plasmodiu	C 487	70	8.0	22137	12	CR954958	CR954958 Danio rer
C 415	71	8.1	25753	4	AF437291	AF437291 Saccharom	C 488	70	8.0	331039	13	AC116979	AC116979 Dictyoste
C 416	71	8.1	129492	4	AC138466	AC138466 Homo sapi	C 489	69.8	8.0	1812	15	AF267199	AF267199 Candidatu
C 417	71	8.1	154561	11	CR391989	CR391989 Zebrafish	C 490	69.8	8.0	13131	2	AX356424	AX356424 Sequence
C 418	71	8.1	155106	5	AC104069	AC104069 Homo sapi	C 491	69.8	8.0	14211	13	AE001368	AE001368 Plasmodiu
C 419	71	8.1	175997	11	BX005417	BX005417 Zebrafish	C 492	69.8	8.0	85916	13	AC117080	AC117080 Dictyoste
C 420	71	8.1	180784	12	CR391963	CR391963 Danio rer	C 493	69.8	8.0	101739	5	AL591400	AL591400 Human DNA
C 421	71	8.1	187665	12	CR450803	CR450803 Danio rer	C 494	69.8	8.0	110000	12	CT009752_3	Continuation (4 of
C 422	71	8.1	208848	12	AC139608	AC139608 Rattus no	C 495	69.8	8.0	110000	13	AE003846_1	Continuation (2 of
C 423	71	8.1	263707	12	CR356223	CR356223 Danio rer	C 496	69.8	8.0	117199	12	CR792440	CR792440 Danio rer
C 424	70.8	8.1	2131	2	CS083950	CS083950 Sequence	C 497	69.8	8.0	145992	5	AC134919	AC134919 Homo sapi
C 425	70.8	8.1	34750	13	AY217738	AY217738 Eimeria t	C 498	69.8	8.0	152359	5	AC093790	AC093790 Homo sapi
C 426	70.8	8.1	110000	12	PFMAL7PL_12	Continuation (13 o	C 499	69.8	8.0	156843	13	AC010838	AC010838 Drosophil
C 427	70.8	8.1	163735	5	AC027277	AC027277 Homo sapi	C 500	69.8	8.0	160032	12	CT027651	CT027651 Danio rer
C 428	70.8	8.1	172816	5	AC093899	AC093899 Homo sapi							
C 429	70.8	8.1	217904	12	CR925736	CR925736 Danio rer							
C 430	70.8	8.1	256698	12	AC158087	AC158087 Bos tauru							
C 431	70.6	8.1	59414	5	AC093112	AC093112 Homo sapi							
C 432	70.6	8.1	63513	12	AC074061	AC074061 Homo sapi							
C 433	70.6	8.1	108908	13	PFMAL3P8	PFMAL3P8 Plasmodiu							
C 434	70.6	8.1	140171	11	BX537131	BX537131 Zebrafish							
C 435	70.6	8.1	146570	13	AC117076	AC117076 Dictyoste							
C 436	70.6	8.1	154564	12	CR848005	CR848005 Danio rer							
C 437	70.6	8.1	157544	11	CR762435	CR762435 Zebrafish							
C 438	70.6	8.1	202069	12	CR388376	CR388376 Danio rer							
C 439	70.6	8.1	220331	12	AC159903	AC159903 Pan trogl							
C 440	70.6	8.1	250029	13	AE014816	AE014816 Plasmodiu							
C 441	70.6	8.1	331039	13	AC116979	AC116979 Dictyoste							
C 442	70.6	8.1	349980	2	AX344555	AX344555 Sequence							
C 443	70.4	8.0	1554	2	CS083940	CS083940 Sequence							
C 444	70.4	8.0	1818	15	AF267223	AF267223 Candidatu							
C 445	70.4	8.0	66993	12	AC138074	AC138074 Homo sapi							
C 446	70.4	8.0	110000	4	CR382132_17	Continuation (18 o							
C 447	70.4	8.0	110000	12	PFMAL13_09	Continuation (10 o							
C 448	70.4	8.0	110000	15	BA000021_0	BA000021 Wiggleswo							
C 449	70.4	8.0	147244	5	AC005969	AC005969 Homo sapi							
C 450	70.4	8.0	170231	11	AL807377	AL807377 Zebrafish							
C 451	70.4	8.0	180450	13	AE014835	AE014835 Plasmodiu							
C 452	70.4	8.0	181322	5	AC027018	AC027018 Homo sapi							
C 453	70.4	8.0	183584	5	AC012492	AC012492 Homo sapi							
C 454	70.4	8.0	217621	12	BX901924	BX901924 Danio rer							
C 455	70.4	8.0	340552	13	PF929354	PF929354 Plasmodiu							
C 456	70.2	8.0	1090	13	AB084761	AB084761 Drosophil							

## ALIGNMENTS

RESULT 1	AX934431	Sequence 8 from Patent WO02074942.	876 bp	DNA	linear	PAT 05-JAN-2004
AX934431	AX934431	Sequence 8 from Patent WO02074942.				
LOCUS	AX934431	Sequence 8 from Patent WO02074942.				
DEFINITION	AX934431	Sequence 8 from Patent WO02074942.				
ACCESSION	AX934431	Sequence 8 from Patent WO02074942.				
VERSION	AX934431.1	GI:40641678				
KEYWORDS						
SOURCE		Campylobacter jejuni				
ORGANISM		Campylobacter jejuni				
REFERENCE		Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.				
AUTHORS		Gilbert, M. and Wakarchuk, W.W.				
TITLE		Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics				
JOURNAL		Patent: WO 02074942-A 8 26-SEP-2002; National Research Council of Canada (CA)				
FEATURES		Location/Qualifiers				
source		1..876				
		/organism="Campylobacter jejuni"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:197"				
CDS		1..876				
		/note="unnamed protein product; Campylobacter"				

alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
jejunii O:19"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04169.1"  
/db\_xref="GI:40641679"  
/translation="MKKVIAGNGPSLKEIDYSLPNDVFRNCQVYFEDKYYLGGK  
CKAVFYTFNPFEEQYTLKHLIQOEYTELIMGNTNQAHLENENFVKTYDFPDA  
HLGVDFPKQLEFNAYFKPHEIYFNQITSGVMCAVAIALGYKEIYLSGIDFYFGDS  
SYAPDTKOENLLKLAPDKNDRSHVIGHSKNTDJKALEFLEKTKYKILYCLCPNSLLA  
NFIELAPNLNSNFIQENKNTKDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKD  
LLRLPSDIKHFKGK"

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231; Indels 0; Gaps 0;  
Matches 876; Conservative 0; Mismatches 0;

Qy 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAATAATGCAAGCAGTGTTTACACCCCTAATTTCTCTTTTGACGCAATAC 180  
Db 121 CTTGGTAAATAATGCAAGCAGTGTTTACACCCCTAATTTCTCTTTTGACGCAATAC 180

Qy 181 ACTTTAAACATTTAAATCAAGATATGAGACCGAATCAATTTATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAAATCAAGATATGAGACCGAATCAATTTATGTTCTTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTTTTCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360

Qy 361 TTTCCAGCAATTTATTCAAATCAAGATTAACCTCAGGGCTCATATGTCGAGTAGCC 420  
Db 361 TTTCCAGCAATTTATTCAAATCAAGATTAACCTCAGGGCTCATATGTCGAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy 541 GATGCTCGCACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600

Qy 601 GAAAAAATTTTAAAAATTTTAAATTTATGCTTCTTAATAGTCTTTTACGCAATTTT 660  
Db 601 GAAAAAATTTTAAAAATTTTAAATTTTATGCTTCTTAATAGTCTTTTACGCAATTTT 660

Qy 661 ATAGAACTAGCGCCCAAAATTTAAATTTTAAATTTTATCATCAAGAAAAAATACTACT 720  
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTTAAATTTTATCATCAAGAAAAAATACTACT 720

Qy 721 AAAGATATCTATCACTTTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780  
Db 721 AAAGATATCTATCACTTTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780

Qy 781 AAAAAAATAAAAAATTAAGAAATGTTTATTTACAGTTGATAAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAAAAATTAAGAAATGTTTATTTACAGTTGATAAAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAAATAA 876

RESULT 2  
BD249794 876 bp DNA linear PAT 17-JUL-2003  
LOCUS Campylobacter glycosyltransferases for biosynthesis of  
DEFINITION gangliosides and ganglioside mimics.  
ACCESSION BD249794.1 GI:33059564  
VERSION BD249794.1  
KEYWORDS JP 2002535992-A/5.  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and ganglioside mimics  
JOURNAL Patent: JP 2002535992-A 5 29-OCT-2002;  
NATIONAL RESEARCH COUNCIL OF CANADA  
COMMENT OS Campylobacter jejuni  
PN JP 2002535992-A/5  
PD 29-OCT-2002  
PF 01-FEB-2000 JP 2000597438  
PR 01-FEB-1999 US 60/118213, 31-JAN-2000 US 06/495406 P1  
PC C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC  
Campylobacter glycosyltransferases for biosynthesis of CC  
gangliosides and  
CC ganglioside mimics

FEATURES  
source  
FH Key Location/Qualifiers  
FT source 1..876  
/organism='Campylobacter jejuni'.  
/mol\_type='genomic DNA'  
/db\_xref='taxon:197'

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTCTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 61 CTACCAATGATTTTGTATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGTATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180  
Db 121 CTTGTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATGAGACCGAATTAATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATGAGACCGAATTAATGTTCTTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATTTTACGATTTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATTTTACGATTTTTCCT 300

Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTAATGCTTTATTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTTAAAGAAATTAATGCTTTATTTAA 360

Qy 361 TTTCCAGCAATTTATTCNAATCAAGATTAACCTCAGGGCTCATATGTCGAGTAGCC 420



```
Db 361 TTTCCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTCAAAAATAAATACTATATGCTTATGCTCCTAATAGTCTTTTACGAAAATTT 660
Db 601 GAAAAAATTTCAAAAATAAATACTATATGCTTATGCTCCTAATAGTCTTTTACGAAAATTT 660
Qy 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCAATCAAGAAATTTTAAATTTTAAATTT 720
Db 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCAATCAAGAAATTTTAAATTTTAAATTT 720
Qy 721 AAAGATATATCTATACCTTTCTAGTGGCTTATGCTTATGCTCCTAATAGTCTTTTAAATTT 780
Db 721 AAAGATATATCTATACCTTTCTAGTGGCTTATGCTTATGCTCCTAATAGTCTTTTAAATTT 780
Qy 781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
Db 781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAATATAA 876
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAATATAA 876

RESULT 3
AR271703 LOCUS AR271703 876 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 8 from patent US 6503744.
ACCESSION AR271703
VERSION AR271703.1 GI:29703248
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6503744-A 8 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;
FEATURES Location/Qualifiers
source 1..876
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 876; DB 2; Length 876;
Best Local Similarity 100.0%; Pred. No. 9.9e-231;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTTCTGCGAAATGGCAACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTTCTGCGAAATGGCAACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Qy 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTGCGTAAAAATCAAGAGCTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
```

```
Db 121 CTGCGTAAAAATCAAGAGCTGTTTACACCCCTAATTTCTTTTGTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCCAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTTAAATTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTTAAATTTTAA 360
Qy 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTTAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTTAAATTTCTA 600
Qy 601 GAAAAAATTTCAAAAATAAATACTATATGCTTATGCTCCTAATAGTCTTTTACGAAAATTT 660
Db 601 GAAAAAATTTCAAAAATAAATACTATATGCTTATGCTCCTAATAGTCTTTTACGAAAATTT 660
Qy 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCAATCAAGAAATTTTAAATTTTAAATTT 720
Db 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCAATCAAGAAATTTTAAATTTTAAATTT 720
Qy 721 AAAGATATATCTATACCTTTCTAGTGGCTTATGCTTATGCTCCTAATAGTCTTTTAAATTT 780
Db 721 AAAGATATATCTATACCTTTCTAGTGGCTTATGCTTATGCTCCTAATAGTCTTTTAAATTT 780
Qy 781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
Db 781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAATATAA 876
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAATATAA 876

RESULT 4
AR481785 LOCUS AR481785 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 8 from patent US 6699705.
ACCESSION AR481785
VERSION AR481785.1 GI:47243420
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6699705-A 8 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES Location/Qualifiers
source 1..876
/mol_type="genomic DNA"
ORIGIN
```

Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTGCTGTAATAAATGCAAAAGCAGTGTTCACACCCCTAAATTTCTTCTTGAGCAATACTAC 180  
Db 121 CTGCTGTAATAAATGCAAAAGCAGTGTTCACACCCCTAAATTTCTTCTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTTAATATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTTAATATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300

Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAAA 360

Qy 361 TTTTCAACCAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC 420  
Db 361 TTTTCAACCAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 GAAAAAATTTTACAAAAATAAAGCAATTTTAAATTTTATCAAGTTGATAAAGATTAAGATTA 840  
Db 481 GAAAAAATTTTACAAAAATAAAGCAATTTTAAATTTTATCAAGTTGATAAAGATTAAGATTA 840

Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 876  
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 876

RESULT 5  
AR527384  
LOCUS AR527384 876 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 8 from patent US 6723545.  
ACCESSION AR527384  
VERSION AR527384.1 GI:53914405  
KEYWORDS

SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 876)  
TITLE Gilbert, M. and Wakarchuk, W. W.  
JOURNAL Polypeptides having .beta.-1,4-GalNAc transferase activity  
Patent: US 6723545-A 8 20-APR-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source Location/Qualifiers  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60

Qy 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTGCTGTAATAAATGCAAAAGCAGTGTTCACACCCCTAAATTTCTTCTTGAGCAATACTAC 180  
Db 121 CTGCTGTAATAAATGCAAAAGCAGTGTTCACACCCCTAAATTTCTTCTTGAGCAATACTAC 180

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTTAATATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTTAATATGTTCTAAT 240

Qy 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300

Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAAA 360

Qy 361 TTTTCAACCAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC 420  
Db 361 TTTTCAACCAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGTATACCAAAATAAAGCAATTTTAAATTTTATCAAGTTGATAAAGATTAAGATTA 840  
Db 481 TCTTATGCTTTTGTATACCAAAATAAAGCAATTTTAAATTTTATCAAGTTGATAAAGATTAAGATTA 840

Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 876  
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 876

Db 781 AAAAAATAAAATTAAGAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTCATATAAGCATTTATTTCAAGGAAATAA 876  
|||||  
Db 841 CCTAGTCATATAAGCATTTATTTCAAGGAAATAA 876  
|||||

RESULT 6

AR609663

LOCUS AR609663 876 bp DNA linear PAT 15-DEC-2004

DEFINITION Sequence 8 from patent US 6825019.

ACCESSION AR609663

VERSION AR609663.1 GI:56664963

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 876)

AUTHORS Gilbert M. and Wakarchuk, W.W.

TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity

JOURNAL Patent: US 6825019-A 8 30-NOV-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES

source Location/Qualifiers

1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAAGTTATTATGCTGGAATGGAACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
|||||

Db 1 ATGAAAAAAGTTATTATGCTGGAATGGAACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
|||||

Qy 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTGGAAGATAATACATAT 120  
|||||

Db 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTGGAAGATAATACATAT 120  
|||||

Qy 121 CTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACATAC 180  
|||||

Db 121 CTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACATAC 180  
|||||

Qy 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAATATGAGCCGAACTTAATATGTTCTTAAT 240  
|||||

Db 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAATATGAGCCGAACTTAATATGTTCTTAAT 240  
|||||

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300  
|||||

Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300  
|||||

Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
|||||

Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
|||||

Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTTCAGGGGTCTATATGTCGAGTAGGCC 420  
|||||

Db 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTTCAGGGGTCTATATGTCGAGTAGGCC 420  
|||||

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTATCAAAATGGGTCA 480  
|||||

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTATCAAAATGGGTCA 480  
|||||

Qy 481 TCTTATGCTTTTGNATCAACACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
|||||

Db 481 TCTTATGCTTTTGNATCAACACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
|||||

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
|||||

Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
|||||

Qy 601 GAAAAAACTTACAAAATAAAACTATATTGCTTATGCTCTTAATAGTCTTTTACCAAATTTT 660  
|||||

Db 601 GAAAAAACTTACAAAATAAAACTATATTGCTTATGCTCTTAATAGTCTTTTACCAAATTTT 660  
|||||

Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 720  
|||||

Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 720  
|||||

Qy 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
|||||

Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
|||||

Qy 781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
|||||

Db 781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
|||||

Qy 841 CCTAGTCATATAAGCATTTATTTCAAGGAAATAA 876  
|||||

Db 841 CCTAGTCATATAAGCATTTATTTCAAGGAAATAA 876  
|||||

RESULT 7

AR689941

LOCUS AR689941 876 bp DNA linear PAT 12-SBP-2005

DEFINITION Sequence 8 from patent US 6905867.

ACCESSION AR689941

VERSION AR689941.1 GI:74471949

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 876)

AUTHORS Gilbert, M. and Wakarchuk, W.W.

TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity

JOURNAL Patent: US 6905867-A 8 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES

source Location/Qualifiers

1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAAGTTATTATGCTGGAATGGAACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
|||||

Db 1 ATGAAAAAAGTTATTATGCTGGAATGGAACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
|||||

Qy 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
|||||

Db 61 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATACTAT 120  
|||||

Qy 121 CTGCGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACATAC 180  
|||||

Db 121 CTGCGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACATAC 180  
|||||

Qy 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAATATGAGCCGAACTTAATATGTTCTTAAT 240  
|||||

Db 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAATATGAGCCGAACTTAATATGTTCTTAAT 240  
|||||

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300  
|||||

Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300  
|||||

Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
|||||

Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
|||||

Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTTCAGGGGTCTATATGTCGAGTAGGCC 420  
|||||

```
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGTCCTATAGTCCTTTTACCAATTTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGTCCTATAGTCCTTTTACCAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATAACTACACT 720
Qy 721 AAAGATATCTCATACACTTCTAGTAGGCTTATGCGAAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTCATACACTTCTAGTAGGCTTATGCGAAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 8
AR691837 AR691837 876 bp DNA linear PAT 13-SEP-2005
LOCUS Sequence 8 from patent US 6911337.
DEFINITION AR691837
ACCESSION AR691837.1 GI:74479548
VERSION
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert M. and Wakarchuk W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAC transferase
JOURNAL Patent: US 6911337-A 8 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES Location/Qualifiers
source 1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 876; DB 2; Length 876;
Best Local Similarity 100.0%; Pred. No. 9,9e-231; Indels 0; Gaps 0;
Matches 876; Conservative 0; Mismatches 0;
Qy 1 ATGAAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Qy 61 CTACCAAAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTGGTAAAAAAGCAAGCAGTGTTTTACACCCCTAATTTCTTCTTTGACCAATACTAC 180
Db 121 CTTGGTAAAAAAGCAAGCAGTGTTTTACACCCCTAATTTCTTCTTTGACCAATACTAC 180
```

```
Qy 181 ACTTTAAAAACATTTTATCCAAATCAAGAATATGAGACCGAACTAAATTTATGTGTCTTAAT 240
Db 181 ACTTTAAAAACATTTTATCCAAATCAAGAATATGAGACCGAACTAAATTTATGTGTCTTAAT 240
Qy 241 TAAACCAAGCTCATCTAGAAAAAGAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCT 300
Db 241 TAAACCAAGCTCATCTAGAAAAAGAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTATTTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATAACTACACT 720
Qy 721 AAAGATATCTCATACACTTCTAGTAGGCTTATGCGAAATTTTCAAAAAATTTTAAATTTT 780
Db 721 AAAGATATCTCATACACTTCTAGTAGGCTTATGCGAAATTTTCAAAAAATTTTAAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 9
AY661458 6047 bp DNA linear BCT 20-JAN-2006
LOCUS Campylobacter jejuni strain CF90-26 truncated putative
DEFINITION glucosyltransferase, putative glycosyltransferase,
beta-1,4-N-acetylglucosaminyltransferase (cgta),
beta-1,3-galactosyltransferase (cgtB), and
alpha-2,3-sialyltransferase (cstII) genes, complete cds; and sialic
acid synthase (neub) gene, partial cds.
VERSION AY661458
KEYWORDS AY661458.1 GI:56159787
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
REFERENCE 1 (bases 1 to 6047)
AUTHORS Koga, M., Gilbert, M., Takahashi, M., Li, J., Koike, S., Hirata, K. and
Yuki, N.
TITLE Comprehensive Analysis of Bacterial Risk Factors for the
Development of Guillain-Barre Syndrome after Campylobacter jejuni
Enteritis
JOURNAL J. Infect. Dis. 193 (4), 547-555 (2006)
```



```

601 GAAAAAATTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Dd |||||||
5471 GAAAAAATTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 5530
Qy |||||||
661 ATAGAATAGCGGCAAAATTTAAATTTAAATTTTATCATACAAAGAAAAATAAATACTACACT 720
Dd |||||||
5531 ATAGAATAGCGGCAAAATTTAAATTTAAATTTTATCATACAAAGAAAAATAAATACTACACT 5590
Qy |||||||
721 AAAGATATATCTCATACCTTCTAGTGAGCGCTTATGGAATAATTTTCAAAAAATAATTAATTTT 780
Dd |||||||
5591 AAAGATATATCTCATACCTTCTAGTGAGCGCTTATGGAATAATTTTCAAAAAATAATTAATTTT 5650
Qy |||||||
781 AAAAAATAAAAAATAAGAAAAATGCTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Dd |||||||
5651 AAAAAATAAAAAATAAGAAAAATGCTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 5710
Qy |||||||
841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876
Dd |||||||
5711 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 5746

RESULT 10
AY422197
LOCUS
DEFINITION
Campylobacter jejuni putative ABC-type transport protein, putative
UDP-glucose-4-epimerase, hypothetical protein, putative
heptosyltransferase I, putative lipid A biosynthesis
acyltransferase, putative two-domain glycosyltransferase, putative
glycosyltransferase, putative
beta-1,4-N-acetylglucosaminyltransferase, putative
beta-1,3-galactosyltransferase, putative
alpha-2,3-sialyltransferase, putative sialic acid synthase,
putative UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine
synthase, putative CMP-NeuSAC synthetase, putative
acetyltransferase, putative glycosyltransferase, putative
heptosyltransferase II, putative phosphoheptose isomerase 1,
putative ADP-heptose synthase, putative
ADP-glycerolmannoheptose-6-epimerase, putative phosphatase, putative
periplasmic cytochrome C, small hypothetical hydrophobic protein,
and putative cation-transporting ATPase genes, complete cds; and
putative transcription termination factor gene, partial cds.
AY422197
AY422197.1 GI:40217896
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
1 (bases 1 to 24425)
Gilbert,M., Godschalk,P.C., Karwaski,M.F., Ang,C.W., Van Belkum,A.,
Li,J., Wakarchuk,W.W. and Endtz,H.P.
Evidence for Acquisition of the Lipopoligosaccharide Biosynthesis
Locus in Campylobacter jejuni GB11, a Strain Isolated from a
Patient with Guillain-Barre Syndrome, by Horizontal Exchange
Infect. Immun. 72 (2), 1162-1165 (2004)
14742567
2 (bases 1 to 24425)
Gilbert,M., Karwaski,M.-F., Cunningham,A. and Brochu,D.
Direct Submission
Submitted (24-SEP-2003) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
source
1. .24425
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="GB11"
/serotype="HS:2"
/db_xref="taxon:197"
complement (21..1715)
/codon_start=1
/transl_table=11
/product="putative ABC-type transport protein"

CDS

```

```

/protein_id="AAR82866.1"
/db_xref="GI:40217897"
/translation="MLKKLFFILSKEDKNFLFLLVFSVFISFIETFAISLVMPFITL
ASDSYPRNKVLSLKEVNLIPVEIIVYFGVGLIVFVFRALLNAVYFPHLLARFSK
GRYHAIAYKFSKFLINVEKFTQKQSEILKSITGEVNLSTMISSEFLLMLFV
LLLYALMLUNIKITLFIIFMLNFIILVKLSPIIKKAGVRRERAMKNFFILNTN
LNNFKFKTKTGKDGVSFLKQASFAFSCANITNESVAAPRIYLESGIFCVLIVFIV
PLVKNESDIIGILSTISIFVLALYRLMPSANRIITSYHDLIYHSLDIYYQLNRQE
BENLEGEKLSNQELIKCNLSFGYEGKKYLPKNLNIKKGKIAIFIGSGCGKSTLV
LIIGLAKPKSGOILIDQELNANNTKNVROKIGVIPONIYLFNDSIAKNIITGDADV
BEKNRVIKQANLEHFIKNLPQGVOTKVGDDGNSLGGQKQRIARALYLEPEMLVL
DEATYALDTQSEAKIMDEIYKISKDKTMIIIAHRLSTIITQCDKVYRLEHGKJKEK"
complement (1709..2695)
/codon_start=1
/transl_table=11
/product="putative UDP-glucose-4-epimerase"
/protein_id="AAR82867.1"
/db_xref="GI:40217898"
/translation="MKILISGGAGYIGSHTLRQFLKTDHEICVLDNLSKGSKIAIEDL
OKIRAFKFEODLSDFQGVKALFEREKDAIVHFAASI EVFESMNPPLKYMMNTVNT
TNLIETCLOTGVNKEIFSTAAATYGEPTPVVSETSP LAPINPYGRSKLMSEVLDDA
SMANPEFKCILRYFNAGACMDYTLGQRYPKATILLIKVAACAAGKRDKLFEGDY
DKDGTGTCRDIHVDDIISAHLAALDYKENSFNVFGYGHGFSVKEVIEAMKKVSG
VDFKVELAPRRAGDPSVLISDASKIRNLTSWQPKYDDELELICKSAFDEWEKQC"
complement (2753..3583)
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAR82868.1"
/db_xref="GI:40217899"
/translation="MILAKLFKKRLKMAKNEGYICVDCESVPDVELIRKTLQFEGSD
LEVSLKALQMOKEQSGSFLPLPYHKIISICAVLSDNFGKFIKVNKIDQNEKEMEN
FNFNTIENYEPKLIVSNGKNFMPVLVLRALKYNLKAATYLDQSDKNYKTFESLKF
HCDLLESUGSNGRGKLDLTCSWGLPGKIDYHGVDEVMKLFTENELEKIEYECSELV
NTYMLFLKYELIKANVSBEDYVDFLSYMRDLFREKKSDRSRSTEVFAKACESEISKVRS"
3613..4641
/codon_start=1
/transl_table=11
/product="putative heptosyltransferase I"
/protein_id="AAR82869.1"
/db_xref="GI:40217900"
/translation="MKIAIVRLSALGDIISAVVLOFTIKNFKKDIEIHWFDKFEGI
LKNPLIDKLYALPLKDKKILKSLKLLKKNYNAVIDLQGLIKSAIVSLSRNN
FGDKNSLKESFAHNFYQKLELDINENFVRLYSUTSFMNLTFDNAKLAFQDVFS
VDENLKQLNNKLDKNEKNLIHVGSVENKNIYPTKTLAILCKLLINEFQOTKWL
AWGNVKEVFAKELVNLGSDIDETHLEAPKNLEBELMAFTKMDMLIIGNDSGPTHLAF
ALNKASITIFGATPSYRNAFQTHINKIIDTGKKIQNAKHIDKSDFCITRIEEDIIFKL
AKGLLNEX"
4631..5518
/codon_start=1
/transl_table=11
/product="putative lipid A biosynthesis acyltransferase"
/protein_id="AAR82870.1"
/db_xref="GI:40217901"
/translation="MKNSDRIYLSLYILKPFVTPMPDCILHFLALIVARIAPHLNKK
HRKINTNLQICFPQYTKERDKLSKIYENFAQFGIDCLOQNTKELINKVNFIN
ENFLIDALAKRPPIIFTAHYKNWEILSLAYAAKYGVISVIGKLSVWYELSSOR
TQFDIELDKKGGIRQMLSKALKERALGILTDQDCVENESVRLKFFKNVYQMGASL
IAORSNALIIPVYAYKEGKFCIEFFKADKSDNASLEULTLYQAQSCSEMIKKRPWEY
FFPHRRPASYNEEYKGA"
5515..6768
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAR82871.1"
/db_xref="GI:40217902"
/translation="MNLKQISVIIIVKNAEQTLLECLNLSKDFDEIILLNNESSDNTL
KIANEFKDFDANLYIYHNAFGFALKNALSYAKNDWILSIDADVLNECEIKELKN
KKIQEDNIIALSRLKNIYKGEWIKACGWMPDYVLRIFNKNFTFNDNLVHESLVLPNSA
LKIYLNGLKHYSYKDIISHLIDKMOYYSMAKQNIHKSGVLRKANLRAFVTFPRNYF
LKNGLFYKYGKFIISVCSALGTFFKMYLYELQROKPKTCALIIITYNOKERLKLVDLS
VKNLALPLNEVLADDSGSKEDTARLIEEYQDFFPCLKHIEWEDEGFKLSKSNKTIK

```

NADSEYIIVIDGMILEKOFIKHELEFAQRKFLQGSRVILNKKESEELNKDORYII  
FNKDPKSSKNFLAKIFYLSKKR  
7064.8233

CDS

```

/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAR82872.1"
/db_xref="GI:40217903"
/translation="MKKIGVVIPIYNEVKYLRCLDSVINQTYTNLEIILVNDGSTD
HSLNIKEYTLKDRITLFDKNGGLSSARNIGIEYFSGYKLNKTKQHIKENSLEIF
QLDGNPNYIKYAKYSQAFNEKDLTNFTYPSIDYIIIFLDSNYKWLKNCIECVIRM
KNVDLWFHDCTYENIKKHKKRMEIPDFKKECIITPKYANRVLGVSGDISFG
WNGMDLFLKQJLKFELINEDIHFGIILPASANKIIVLSOKLYLCLRLANSISN
HDKKIKANVSEFKDIYETFGENAKENYLAASRVITALKLIEFPKQKQENALA
IKETFLPCYAKALMIKKFKDPLNLKEQVLVLIKPFITQTLKLPYDIWKFWOKIKNI"
8303.9346
/codon_start=1
/transl_table=11
/product="putative

```

CDS

```

beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAR82873.1"
/db_xref="GI:40217904"
/translation="MLFQSYFVKIICLIFPRKIRHKIKKTFLLKNIQORDKIDSYLPK
KTLVOIKYNNEDILINKAI IREHGKVENYDEKSDPKPSPLNPAFIRVKNEAITL
KASLESLPALQGVIGYNDCTDGSSEILILEFCQYPSFTPIKYPYEQIQNPKSEEN
KLISYNYVASFIPKDEWLKIDVDHIYDAKLYKSPYIPKPKYDVVSYSRVDIHPFN
DNFPLCKONNDILKBPDCCLINNNYLKWEVLIDRINNKKATQSPSSNIHSLE
QLYKRIILPHTLNNYHFPFLKRAQDIYKYNWISIEEFKPYLQNIHKKIEPSSI
SKETLKKIFLTLF"
complement(9332..10237)
/codon_start=1
/transl_table=11
/product="putative beta-1,3-galactosyltransferase"
/protein_id="AAR82874.1"
/db_xref="GI:40217905"
/translation="MPKISILPTVYNEQVIARAIESCINOTFKDIEIIVVDGNDN
SINIAEYKSKDKRIKIIHNEKLLGLLRAEYGVKVANSYIMFLDPPDDHLELNACEE
CIIKLEQDEVDLFFNAIVESNVIYKDFNSGFYSKKEFEVKKIYAKNLYWTMG
KLIRKLYEAPLALSKEDVKNMAEDVLLYPMLSQAKIAYMNCNLVHYVNNNS
ICNTKNEVLVNNIIOELQVLNLYLRQNYILNKYCSVLVLYIKLYIKIKRTKLM
VTLAKINILTLKILPKYKFLKQC"
10326..11201
/codon_start=1
/transl_table=11
/product="putative alpha-2,3-sialyltransferase"
/protein_id="AAR82875.1"
/db_xref="GI:40217906"
/translation="MKKVIAGNGPSLKEIDYSLPNDPVDVFCNQYFEDKYVLGKK
CKAVYTPNFFEQYTLKHLIQNOYETELIMCSYNNQAHLENENFVKTFYDPDA
HLGVDFPKQENLNAEYKFEHIEYFNQRTISGVYMCVAIAIAGYKEIYLSGIDIFYGSA
SYAFDTKQENLLKLPDNDKSHRYHGSKNITDIKALEFLEKTKYIKLYCLCPNSLLA
NFTELAPLNSNFTIOEKNNYTIOLILIPSESEAYKFKSNINFKIKIKENYVYKLKD
LLRLPSDIKHVFKGK"
11202..12242
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAR82876.1"
/db_xref="GI:40217907"
/translation="MKKIKIQNIILSEKAPLVPEIGINHSLELAKIMVDAFSA
GAKIIRHQTHIVDEMSKAARKVIPGNAKISYIEMOKCALDYKDELALKEYETKGL
VYLSTPFSRAGARLEMDGVSFAKIGSGECNNYPLIKHIAAFKPKMIVSTGMNIESI
KPTVLLDNEIIPVLMHTNLYPTPHNLNMLKLEKFSKCMVGLSDHTNLAAC
LGAVALGACVLERHTDSMDRSGDPIVCSMDTKALKELIMQSEOMATMRGNESKKA
KQEQVTDIFAFASVUSIKDIKKEVLSMDNIWVRPGLGISAVEFENILGKALRDI
ENOTQLSYEDFA"
12239..13357
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine
2-epimerase/N-acetylmannosamine synthase"
/protein_id="AAR82877.1"
/db_xref="GI:40217908"

```

CDS

```

/codon_start=1
/transl_table=11
/product="putative alpha-2,3-sialyltransferase"
/protein_id="AAR82875.1"
/db_xref="GI:40217906"
/translation="MKKVIAGNGPSLKEIDYSLPNDPVDVFCNQYFEDKYVLGKK
CKAVYTPNFFEQYTLKHLIQNOYETELIMCSYNNQAHLENENFVKTFYDPDA
HLGVDFPKQENLNAEYKFEHIEYFNQRTISGVYMCVAIAIAGYKEIYLSGIDIFYGSA
SYAFDTKQENLLKLPDNDKSHRYHGSKNITDIKALEFLEKTKYIKLYCLCPNSLLA
NFTELAPLNSNFTIOEKNNYTIOLILIPSESEAYKFKSNINFKIKIKENYVYKLKD
LLRLPSDIKHVFKGK"
11202..12242
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAR82876.1"
/db_xref="GI:40217907"
/translation="MKKIKIQNIILSEKAPLVPEIGINHSLELAKIMVDAFSA
GAKIIRHQTHIVDEMSKAARKVIPGNAKISYIEMOKCALDYKDELALKEYETKGL
VYLSTPFSRAGARLEMDGVSFAKIGSGECNNYPLIKHIAAFKPKMIVSTGMNIESI
KPTVLLDNEIIPVLMHTNLYPTPHNLNMLKLEKFSKCMVGLSDHTNLAAC
LGAVALGACVLERHTDSMDRSGDPIVCSMDTKALKELIMQSEOMATMRGNESKKA
KQEQVTDIFAFASVUSIKDIKKEVLSMDNIWVRPGLGISAVEFENILGKALRDI
ENOTQLSYEDFA"
12239..13357
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine
2-epimerase/N-acetylmannosamine synthase"
/protein_id="AAR82877.1"
/db_xref="GI:40217908"

```

CDS

```

/codon_start=1
/transl_table=11
/product="putative alpha-2,3-sialyltransferase"
/protein_id="AAR82875.1"
/db_xref="GI:40217906"
/translation="MKKVIAGNGPSLKEIDYSLPNDPVDVFCNQYFEDKYVLGKK
CKAVYTPNFFEQYTLKHLIQNOYETELIMCSYNNQAHLENENFVKTFYDPDA
HLGVDFPKQENLNAEYKFEHIEYFNQRTISGVYMCVAIAIAGYKEIYLSGIDIFYGSA
SYAFDTKQENLLKLPDNDKSHRYHGSKNITDIKALEFLEKTKYIKLYCLCPNSLLA
NFTELAPLNSNFTIOEKNNYTIOLILIPSESEAYKFKSNINFKIKIKENYVYKLKD
LLRLPSDIKHVFKGK"
11202..12242
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAR82876.1"
/db_xref="GI:40217907"
/translation="MKKIKIQNIILSEKAPLVPEIGINHSLELAKIMVDAFSA
GAKIIRHQTHIVDEMSKAARKVIPGNAKISYIEMOKCALDYKDELALKEYETKGL
VYLSTPFSRAGARLEMDGVSFAKIGSGECNNYPLIKHIAAFKPKMIVSTGMNIESI
KPTVLLDNEIIPVLMHTNLYPTPHNLNMLKLEKFSKCMVGLSDHTNLAAC
LGAVALGACVLERHTDSMDRSGDPIVCSMDTKALKELIMQSEOMATMRGNESKKA
KQEQVTDIFAFASVUSIKDIKKEVLSMDNIWVRPGLGISAVEFENILGKALRDI
ENOTQLSYEDFA"
12239..13357
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine
2-epimerase/N-acetylmannosamine synthase"
/protein_id="AAR82877.1"
/db_xref="GI:40217908"

```

CDS

```

/codon_start=1
/transl_table=11
/product="putative alpha-2,3-sialyltransferase"
/protein_id="AAR82875.1"
/db_xref="GI:40217906"
/translation="MKKVIAGNGPSLKEIDYSLPNDPVDVFCNQYFEDKYVLGKK
CKAVYTPNFFEQYTLKHLIQNOYETELIMCSYNNQAHLENENFVKTFYDPDA
HLGVDFPKQENLNAEYKFEHIEYFNQRTISGVYMCVAIAIAGYKEIYLSGIDIFYGSA
SYAFDTKQENLLKLPDNDKSHRYHGSKNITDIKALEFLEKTKYIKLYCLCPNSLLA
NFTELAPLNSNFTIOEKNNYTIOLILIPSESEAYKFKSNINFKIKIKENYVYKLKD
LLRLPSDIKHVFKGK"
11202..12242
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAR82876.1"
/db_xref="GI:40217907"
/translation="MKKIKIQNIILSEKAPLVPEIGINHSLELAKIMVDAFSA
GAKIIRHQTHIVDEMSKAARKVIPGNAKISYIEMOKCALDYKDELALKEYETKGL
VYLSTPFSRAGARLEMDGVSFAKIGSGECNNYPLIKHIAAFKPKMIVSTGMNIESI
KPTVLLDNEIIPVLMHTNLYPTPHNLNMLKLEKFSKCMVGLSDHTNLAAC
LGAVALGACVLERHTDSMDRSGDPIVCSMDTKALKELIMQSEOMATMRGNESKKA
KQEQVTDIFAFASVUSIKDIKKEVLSMDNIWVRPGLGISAVEFENILGKALRDI
ENOTQLSYEDFA"
12239..13357
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine
2-epimerase/N-acetylmannosamine synthase"
/protein_id="AAR82877.1"
/db_xref="GI:40217908"

```

CDS

```

/codon_start=1
/transl_table=11
/product="putative alpha-2,3-sialyltransferase"
/protein_id="AAR82875.1"
/db_xref="GI:40217906"
/translation="MKKVIAGNGPSLKEIDYSLPNDPVDVFCNQYFEDKYVLGKK
CKAVYTPNFFEQYTLKHLIQNOYETELIMCSYNNQAHLENENFVKTFYDPDA
HLGVDFPKQENLNAEYKFEHIEYFNQRTISGVYMCVAIAIAGYKEIYLSGIDIFYGSA
SYAFDTKQENLLKLPDNDKSHRYHGSKNITDIKALEFLEKTKYIKLYCLCPNSLLA
NFTELAPLNSNFTIOEKNNYTIOLILIPSESEAYKFKSNINFKIKIKENYVYKLKD
LLRLPSDIKHVFKGK"
11202..12242
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAR82876.1"
/db_xref="GI:40217907"
/translation="MKKIKIQNIILSEKAPLVPEIGINHSLELAKIMVDAFSA
GAKIIRHQTHIVDEMSKAARKVIPGNAKISYIEMOKCALDYKDELALKEYETKGL
VYLSTPFSRAGARLEMDGVSFAKIGSGECNNYPLIKHIAAFKPKMIVSTGMNIESI
KPTVLLDNEIIPVLMHTNLYPTPHNLNMLKLEKFSKCMVGLSDHTNLAAC
LGAVALGACVLERHTDSMDRSGDPIVCSMDTKALKELIMQSEOMATMRGNESKKA
KQEQVTDIFAFASVUSIKDIKKEVLSMDNIWVRPGLGISAVEFENILGKALRDI
ENOTQLSYEDFA"
12239..13357
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine
2-epimerase/N-acetylmannosamine synthase"
/protein_id="AAR82877.1"
/db_xref="GI:40217908"

```

```

/translation="MKKILFITGRADYSKISKLMYRVQNSSEFELYIFATGMHLNKN
FGYTVKELYKNGPKNIYEFINYDVKYOTDKALATTIDGFSRYVLANEKDOLI VVHGDR I
EPLAAAIYGALNNILVAHISGGEISGTFIDDSIRHAI SKLAHIHLANDFEKRLMQLG
EDEKSIIFIGSPDELLDNKINSEAKKYDINYENYVALLMFHPVTTEITSIKNQAD

Query Match      100.0%; Score 876; DB 15; Length 24425;
Best Local Similarity 100.0%; Pred. No. 2.4e-231;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60
Db 10326 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 10385
Qy 61 CTACCAATGATTTTGTATGTTAGATGTAATCAATTTTATTTTGAAGATAAACTACTAT 120
Db 10386 CTACCAATGATTTTGTATGTTAGATGTAATCAATTTTATTTTGAAGATAAACTACTAT 10445
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Db 10446 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 10505
Qy 181 ACTTTAAACATTTAATCCAAAATCAGAATATGAGACCGAAGTAAATATGTTGTTCTAAT 240
Db 10506 ACTTTAAACATTTAATCCAAAATCAGAATATGAGACCGAAGTAAATATGTTGTTCTAAT 10565
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATTATTTCCT 300
Db 10566 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATTATTTCCT 10625
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 10626 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 10685
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTTATATGTTGTCAGTAGCC 420
Db 10686 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTTATATGTTGTCAGTAGCC 10745
Qy 421 ATAGCCCTAGCATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGGTCA 480
Db 10746 ATAGCCCTAGCATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGGTCA 10805
Qy 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 10806 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 10865
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600
Db 10866 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 10925
Qy 601 GAAAAAATTTACAAAATAAAAATATATATGCTTATGCTTCTTAAGTCTTTTACCAATTTT 660
Db 10926 GAAAAAATTTACAAAATAAAAATATATATGCTTATGCTTCTTAAGTCTTTTACCAATTTT 10985
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTCAATTTTATCATACAGAAAAAATAACTACACT 720
Db 10986 ATAGAACTAGCGCCAAATTTTAAATTTTCAATTTTATCATACAGAAAAAATAACTACACT 11045
Qy 721 AAAGATATACTCATACCTTCTTAGTGAGGCTTTATGAAAAATTTTCAAAAAATATTAATTT 780
Db 11046 AAAGATATACTCATACCTTCTTAGTGAGGCTTTATGAAAAATTTTCAAAAAATATTAATTT 11105
Qy 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 11106 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 11165
Qy 841 CCTAGTGATATAAGCATTATTTCAGAGAAAAATAA 876
Db 11166 CCTAGTGATATAAGCATTATTTCAGAGAAAAATAA 11201

RESULT 11
AF167344 AF167344 24437 bp DNA linear BCT 23-JAN-2004
LOCUS
```



DEFINITION Campylobacter jejuni HS:19 lipooligosaccharide biosynthesis locus, partial sequence.

ACCESSION AF167344

VERSION AF167344.2 GI:41058381

KEYWORDS

SOURCE

ORGANISM Campylobacter jejuni

REFERENCE

AUTHORS Gilbert,M., Brisson,J.R., Karwaski,M.F., Michniewicz,J., Cunningham,A.M., Wu,Y., Young,N.M. and Wakarchuk,W.W.

TITLE Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384. Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-mhz (1)h and (13)C NMR analysis

JOURNAL J. Biol. Chem. 275 (6), 3896-3906 (2000)

PUBMED 10660542

REFERENCE

AUTHORS Gilbert,M., Cunningham,A.-M., Karwaski,M.-F., Michniewicz,J., Wu,Y., Young,N.M. and Wakarchuk,W.W.

TITLE Direct Submission

JOURNAL Submitted (08-JUL-1999) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada

REFERENCE

AUTHORS Gilbert,M.

TITLE Direct Submission

JOURNAL Submitted (22-JAN-2004) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada

REMARK Nucleotide and amino acid sequence updated by submitter

COMMENT On Jan 22, 2004 this sequence version replaced gi:6969981.

FEATURES

source

1..24437

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strain="ATCC 43446"

/serotype="HS:19"

/db\_xref="ATCC:43446"

/db\_xref="taxon:197"

complement(21..1715)

/codon\_start=1

/transl\_table=11

/product="putative ABC-type transport protein"

/protein\_id="AAR399161.1"

/db\_xref="GI:41058384"

/translation="MLKLFILSKEDKNFLFLLVFSVFSPIETFAISLWMPFITL ASDPSVDRNKYLISLKEYLNIPVEIIVYFVGLIVFYVFRALLNAYYPHLLARFSK GRHTAAYKFSFLNINYEFTQKNQSEILKSTIGVGLIVFYVFRALLNAYYPHLLARFSK LLLYALMLLYNITPLSLIPVNLNAPFILVKILSPITIKKAGLRREEMKPNFEILNTN LNNFKITLKTEDGVLSLPAQSEAFSKANITNESVAAPRIYLEGIFGLVFIYV FLVKNESDISGILSTISFVLAIRLMPNANRITISYHDLIYVHSLIDIYQNLROE EENLGEKLSFNQELKICNLSFGYEGKHLFKNLNLNIKKGEKIAFISGCGKSTLV DLIIGLLKPEGQILDKQELNANAKNQKIGYIPQNIYLFNDSIAKNTIFGDADV EEKLSVKIQANLEHFTKNLPQGVQTKVGGGNSLGGQKQRIARALAYLEPILVL DEATSLDQSEAKINDEIYKISKDKTMIHIAHRLSTITQCDKYRLEHGLKKEEK"

complement(1709..2695)

/codon\_start=1

/transl\_table=11

/product="putative UDP-glucose-4-epimerase"

/protein\_id="AAR399162.1"

/db\_xref="GI:41058385"

/translation="MKLISGGAGYIGSHTRQLQKTDHETICVLDNLSKGSKIAIEDL QKTRAFKPRQDLSDFQGVKALPEREKFDALVHPAASTIEVFSMNPVKYVINNTVT LNLIECMQGVNKEIFSSVTAATYGEQTPVSETSPALPNPGRSKLMSEVLRDA SWANPFKICLRIFNVAGACMDYLCQRPKATLLKVAABCAAGRKDLKIFPGDYI TDKDGTCIRDTIHVDDISSAHLAALDYQENESVFNNGVGHGFSVEKVEAMKVVSG VDFPKLAPRRAGDPSVLIISDAKIRNLTSWQPKYDDLEILICKSAFWEKQC"

complement(2753..3583)

/inference="non-experimental evidence, no additional details recorded"

/codon\_start=1

CDS

3613..4641

/note="formerly ORF-01; similar to Helicobacter pylori RfaC"

/transl\_table=11

/codon\_start=1

/product="heptosyltransferase I"

/protein\_id="AAF34138.2"

/db\_xref="GI:41058382"

/translation="MKIAIVRLSALGDIIQSAVVLQFIKFKKIDBIHMFVDEKFEGI LKNHPLIDKLYALPLDKILKSLKILKARKNNYNAVIDLQGLIKSAIVSRILSRNN FGDKNSLKESFAHNFYKQLELDYNNVYVRLSITSPMLNTDFNAKLAPQDVFS VDENIKQLNNKLIKDKNEKNLIIHVGSSVENKIYPTKLAILCKLLINEFOOTKLWL AGNVNKEYEFAKEVLSLGSIDETHIELAPKFNLEELMAFTKMDLLIIGNDSGTHLAF ALNKASITIFGATPSYNAFQTHINKIIDTGKIQNAKHIDKSDFCITRIEEDIFKL AKGLINEX"

4631..5518

/note="formerly ORF-02; similar to Helicobacter pylori WaaM"

/transl\_table=11

/product="acyltransferase"

/protein\_id="AAF34139.1"

/db\_xref="GI:6969984"

/translation="MKNSDRIYLSLYILKFFVTFMPDCILHFLALIVARIAFLNKK HRKIINTNLQICFPYQTKERDKLSKIYENFAQFIDCLQNTTKKILKNVFNIN ENFLDALALKRPIIFTAHYGNWEILSLAYAGAI SIVGKSLVYMYEILSOSR TQFDIELDKKGGIROMLSALKKRALGILTQDCVENESVRLKPFNKNVYOMGASL IAORSNALIIPVAYKEGKFCIEFFKAKDSQNASLEELTYQAQSCSEMIKRPWEY FFFHRRFASNYEEIYKAK"

5515..6771

/note="formerly ORF-03; similar to Neisseria meningitidis LgtF"

/codon\_start=1

/transl\_table=11

/product="glycosyltransferase"

/protein\_id="AAF34140.1"

/db\_xref="GI:6969985"

/translation="MNLKQISVIIIVKNAEQTLLECLNSLKDPEDEILLNNSSDNTL KIANEFKDFANLYIYHNAFIGFAGLKNLALSYAKNDWILSIDADEVLNECIELKQN LKLOEDNIALSRKNLYKEGWIKAGWPDYVLRIPKNPFRPNDNLVHESLVLPSNA KKIYLNGLKHYSYKDI SHLIDKMQYISLWAKONHKHSGVLKANLRAFWTFRNFT LKNGFLYKGFIIISVCSALGTFYKMYLQYORQKPTCALIITYNOKERLUKVLVD SVKNLAFIPNEVLADGSKEDTARLIEYQKDFPCPLKHIMQEDSGFLKSGRNKTI KNADEYIIVIDGMILEKDFI KEHLEPAQKRLFQGSRVILNKESEELINKDEYRI IFNKDSKSKNSFLAKIFYLSLKR"

7067..8236

/note="formerly ORF-04; similar to Streptococcus pneumoniae Cps14J"

/codon\_start=1

/transl\_table=11

/product="glycosyltransferase"

/protein\_id="AAF34141.1"

/db\_xref="GI:6969986"

/translation="MKKIGVVIPIYVNEKYRECLDSVINQTYTNLEIILVNDGSTDE HSLNTAKPYTLDKRITLFDKNGGLSSARNIGIEYFSGEYKLNKTKQIKENSLEIF QLDGNPNYIKAYKSSOANNEKDLTNFTYPSIDVIIFLDSNLYWKLNCIECVIRM KNDVLFWDHCTYEDNINKKHKKRMELPDFKKECIIITPEYANLVSGRSDISFG WGMIDNFLKQIKLKFINFINEIDHFGIILFASANKIYVLSQKLYLCRLRANSISN HNKITKANVSEYFKDIYETFGENAKENYLAASRYTALKLIEFFQKQKVENALA IKETFLPCYAKKALMIKKFKDPLNKEQLVLIKPFIQTKLPDYIMKFWQIKNI"

8306..9349

/gene="cgtA"

8306..9349

gene

CDS







```

RESULT 13
AF401528      12390 bp      DNA      linear      BCT 03-JAN-2002
LOCUS
DEFINITION
Campylobacter jejuni O:36 heptosyltransferase I (waaC) gene,
partial cds; lipid A biosynthesis acyltransferase (htrB), putative
two-domain glycosyltransferase, putative glycosyltransferase,
truncated beta-1,4-N-acetylglactosaminyltransferase (cgta-I),
beta-1,3-galactosyltransferase (cgtB), alpha-2,3-sialyltransferase
(cst-II), sialic acid synthase (neuB), putative
N-acetylglucosamine-6-phosphate 2-epimerase (neuCl),
beta-1,4-N-acetylglactosaminyltransferase (cgta-II), CMP-Neu5Ac
synthetase (neuA), putative acetyltransferase, and putative
glycosyltransferase (waaF) genes, complete cds; and
heptosyltransferase II (waaP) gene, partial cds.
ACCESSION
AF401528
VERSION
AF401528.1 GI:15718482
KEYWORDS
Campylobacter jejuni
SOURCE
Campylobacter jejuni
ORGANISM
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 12390)
Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E.,
Michniewicz,J., Cunningham,A.M. and Wakarchuk,W.W.
The Genetic Bases for the Variation in the Lipid-oligosaccharide of
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF
SIALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
J. Biol. Chem. 277 (1), 327-337 (2002)
11689567
JOURNAL
PUBMED
REFERENCE
2 (bases 1 to 12390)
Gilbert,M., Michniewicz,J., Karwaski,M.-F., Cunningham,A. and
Wakarchuk,W.W.
Direct Submission
Submitted (22-JUL-2001) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
Location/Qualifiers
1..12390
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43456"
/serotype="O:36"
/db_xref="ATCC:43456"
<1..348
/db_xref="taxon:197"
<1..348
/gene="waaC"
<1..348
/gene="waaC"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAL05984.1"
/db_xref="GI:15718483"
/translation="FAKEVLNLSGDETHIELAPKNLELMAFTKMDLLIIGNDSP
THLAPALMKASTITIGATPSYRNAFOTHIKINDTGKKIQNAKHDKDSFCITRIEE
DIFKLAKGLLNK"
338..1225
/gene="htrB"
338..1225
/gene="htrB"
/codon_start=1
/transl_table=11
/product="lipid A biosynthesis acyltransferase"
/protein_id="AAL05985.1"
/db_xref="GI:15718484"
/translation="MKNSDRIVLSYILKFFVTMPDCILHFLALIVARIAPHLNKK
HRKINTNQICFPQITQERDKLSUKIYENFAQGDICIQNTTKKILNKVFNIN
ENFLDALAKLPGIITFTAHYGNWEILSLAYAKGAIISVKGKLSKYEYVQMGASL
TQFDIELDKKGIRQMLSALKERALGILTDQCVENESVRLKFFKNFVNVMQASL
IAQRNALIIPVYAKGGGKCTIEFFKAKDSQNSDEELTLQAQSCSEMIKKRPWEY
FFEHRRFASYNIEIYKAK"
1222..2769

```

```

/note="orf3"
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAL05986.1"
/db_xref="GI:15718485"
/translation="MNLKQISVIIIVKNAEQTLLECLNSLKDFDEIILLNNESIDNTL
NIANEPKQDPANLIYVHNAIFGALKNLALSIAKNDWILSIDADEVLENECICKKN
LEQEDNIIALSRNLYKGWIKACGWSPDVLRIFNKNFTRFNDNLVHESLVLPSNV
KTHLKNGLKHVAFRDIHSLINKMOYSSILWAKONIHKKSGVLKANLFAWTFERNYF
LKMFLYGYGFTISVNCALGTFKYMKLVELQKQKTKCALIITYNQKERLKLVD
SVNKLFLPNEVLIAADGSKEDTARLIEEYOKDPPCSLKHHLWQDEDEGFAAKNNNAI
RVKSGYIILIDGMILEKDFIKNHFEFAKRVLOQSGRVILNKNESEEILKNDDYSL
AFSKDPKYSNDFLAKIYIKSIDQKFFNTKEIKGSKTCNNSFFKADFDELDFGN
ENFIGMGREDSFVARFLPNKGFPRRLKFAIAHYIHEENSXKMLESHQIYLDTIK
NKKLSWR"
2760..3944
/note="orf4"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAL05987.1"
/db_xref="GI:15718486"
/translation="MEIKHEENRCSYNIQRRRIFERMFNSVINQTYTNLEILVNDG
STDESLNIAKEYTLKDERFILFDKNGGLSSARNVGIEYFSGEYKLNKQTHIKENS
LIEFOLDGNPNYNIYKAYSSQAENNEKDLTNFTYDNIDYIIFLSDDDYBELNCTIEEC
VIRKMVDVLMFDHDCCTYEDNIKNKHKTMEIFDFKKECIITFKEDYANRALSGSRD
ISFGWGMIDFNLKLIKLPINFLINEDIHFGIILPASANKIYLSOKLYLCRLAN
SISNHRKKTIPKANVSEYFKDIYETFGENAKEKNYLKAASRVTTALKLIEFFPKQKSE
NALAIKETFLPCYAKKALMTKKPKDPLNLKEQLVLKPFITQTLFYDIWFKWQKIKK
IN"
3988..4248
/gene="cgtA-I"
3988..4248
/gene="cgtA-I"
/note="orf5-I; premature translation stop due to a one bp
deletion"
/codon_start=1
/transl_table=11
/product="truncated
beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAL05988.1"
/db_xref="GI:15718487"
/translation="MLFQSYFVKIICLIFIPRKIRHKIKKTFLLKNIQDKIDSYLPK
KILIQKNYNNEDLIKLNKAITGGIKDILLIMKNLKIQNLL"
complement(5016..5921)
/gene="cgtB"
complement(5016..5921)
/gene="cgtB"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAL05989.1"
/db_xref="GI:15718488"
/translation="MPKISILPTYNVEQVIARAIESCINOTFKDIEIIVVDDCGNDN
SINIAKEYSKDKRIKTIHNEKNLGLLARVEGVKVANSPIYIMFLDPDDYLELACEE
CIKILDEQDEVDLVFFNAIVESNVISYKDFDFNSGFSKKEFVKIIAKNLYWTWVG
KLIRKKLYLEAFSLRLEKDVKNMAEDVLIIYPMLSQAQKIATYNNCLYHVPNNNS
ICNTKNEVLVNNNIQELQVLNLRQNYILNKYCSVLVLYLLIYIYIYIKIRTKLM
VTLAKINILTLKILFKYKFLKQC"
6010..6885
/gene="cst-II"
6010..6885
/gene="cst-II"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAL05990.1"
/db_xref="GI:15718489"

```

NFIELAPNLNSNFIIOEKNNYTKDILIPSEAYKFKSNINPKKIKENNYVYKLIKD  
LRLPDSIDKHYPKKGK"  
6886. .7926  
/gene="neuB1"  
6886. .7926  
/genes="neuB1"  
/codon\_start=1  
/transl\_table=11  
/product="sialic acid synthase"  
/protein\_id="AAL05991.1"  
/db\_xref="GI:15718490"  
/translation="MKELIQNIIIIEEKAPLVVPEIGINHNSLELAKIMVDAFSA  
GAKIIKHQTHIVDEDSKAAKVPIGNAKISYIEIMQKCALDYKDELALKEYTEKLSL  
VYLSTPFRAGANRLDMGVSAFKIGSGECNNYPLKIHIAAFKPMIVSTGMSIESI  
KPTVKILLONEIPVLMTNLYPTPHNLVRLNMLELKEFSCMVGSLSDHTTNDNLAC  
LGAVALGACVLERHFDSDHRSRSPDIVCSMDTKALKELIIQSEOMAIMKGNESKRA  
KOEQVTDIDFAFASVWSIKDIKKGEVLSMDNIWVRPGLGGISAAEFENILGKALRDI  
ENDTQLSYEDFA"  
7923. .9047  
/gene="neuC1"  
7923. .9047  
/genes="neuC1"  
/codon\_start=1  
/transl\_table=11  
/product="putative N-acetylglucosamine-6-phosphate  
2-epimerase"  
/protein\_id="AAL05992.1"  
/db\_xref="GI:15718491"  
/translation="MKKILFITRADYSKISLMYRVQNSSEFELYIFATGMHLSKN  
FGYTVKELYKNGKNIYEFINYQYOTDKALATTIDGFSRYVNEIKPDLIVVHGDR  
EPLAAIVGALNNILVAHIEGGEISGTIDSLRHAI SKLAHILHVNDEPAKRRLMOLG  
EBEKSIFIGSPLELLNDNKISLSEAKKYDINENYALLMHPHVPTEITSIKNOAD  
NLVKAIQSNKNIIVIPNNDLGFEILQSEELKNNPRFKLPFSURFEYFETILLKNA  
DFTIGNSCILKALYLTAGILVGRONGRLNENTLVNNSDEILKAINTIHKQ  
DLFSTKLEILDLSGFEVLLSGFEFKLSQKVFKDIKHA"  
9040. .9987  
/gene="cgtA-II"  
9040. .9987  
/genes="cgtA-II"  
/notes="orf5-II"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,4-N-acetylglucosaminyltransferase"  
/protein\_id="AAL05993.1"  
/db\_xref="GI:15718492"  
/translation="MLKKIISLYKYSISKSLVLDNEHFTIKENKNIYKHKHGFDFD  
ERAKDVKPLNPGFIRVKNELTLRVLSLEILPALQORHIIAYNDCCDGSSEELILEPC  
KQYPNFIACKYPVKVLEPNKNEENKLYSVNWAASFIPLDEWFIKIDVDHYVDKKL  
YKSFYRIDENKALCYPRINFIILNGNIYVQNSGNYGFIGGGQDILLIKRNSSFIER  
VSKSQWIDPKGDIIEBLYSEQVLSQVSKILQAPLIQWHPALKYRRNDYQQYLDLILS  
LEEFQAFHRKSKEARKIDFAMLRPVIEQILKRFQGEIK"  
9984. .10649  
/gene="neuA1"  
9984. .10649  
/genes="neuA1"  
/codon\_start=1  
/transl\_table=11

Query Match 99.6%; Score 872.8; DB 15; Length 12390;  
Best Local Similarity 99.8%; Pred. No. 2.4e-230;  
Matches 874; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAGG 60  
DB 6010 ATGAAAAAGTTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAGG 6069  
QY 61 CTACCAAAATGATTGATGATTAGATGTAATCAATTTTATTTGAAGATAATATCTAT 120  
DB 6070 CTACCAAAATGATTGATGATTAGATGTAATCAATTTTATTTGAAGATAATATCTAT 6129  
QY 121 CTTGGTAAAAAAGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGGACGATATCTAC 180  
DB 6130 CTTGGTAAAAAAGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGGACGATATCTAC 6189

QY 181 ACTTTAAAAACATTAAATCCAAATCAAGAATATGAGCGAACTAATTAATGTGTTCTAAT 240  
DB 6190 ACTTTAAAAACATTAAATCCAAATCAAGAATATGAGCGAACTAATTAATGTGTTCTAAT 6249  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCCCT 300  
DB 6250 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCCCT 6309  
QY 301 GATCCTCATTTGGGATATGATTTTTTAAACCAACCTTAAAGAAATTAATCTCTTATTTTAAA 360  
DB 6310 GATCCTCATTTGGGATATGATTTTTTAAACCAACCTTAAAGAAATTAATCTCTTATTTAAA 6369  
QY 361 TTTCCAGAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTGTGCGAGTAGCC 420  
DB 6370 TTTCCAGAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATATGTGTGCGAGTAGCC 6429  
QY 421 ATAGCCCTTAGGATACAAAAATTTATCTTTTCGGGAAATGATTTTTATCAAAATGGGTCA 480  
DB 6430 ATAGCCCTTAGGATACAAAAATTTATCTTTTCGGGAAATGATTTTTATCAAAATGGGTCA 6489  
QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 6490 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 6549  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600  
DB 6550 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 6609  
QY 601 GAAAAAACTTACAAAAATAAACTATATTGCTTTATCTCTTAATAGTCTTTTAGCAAAATTTT 660  
DB 6610 GAAAAAACTTACAAAAATAAACTATATTGCTTTATCTCTTAATAGTCTTTTAGCAAAATTTT 6669  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATAACAAGAAAAATAACTACACT 720  
DB 6670 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATAACAAGAAAAATAACTACACT 6729  
QY 721 AAAGATATATCATACCTTCTAGTCAGGCTTATGCAAAATTTTCAAAAAATTAATTTT 780  
DB 6730 AAAGATATATCATACCTTCTAGTCAGGCTTATGCAAAATTTTCAAAAAATTAATTTT 6789  
QY 781 AAAAAATATAAAATTTAAAGAAATCTTTATTACAAAGTTGATAAAAAGATCTATTAAAGATTA 840  
DB 6790 AAAAAATATAAAATTTAAAGAAATCTTTATTACAAAGTTGATAAAAAGATCTATTAAAGATTA 6849  
QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAAATAA 876  
DB 6850 CCTAGTGATATAAGCATTATTTCAAAGGAAAAATAA 6885

RESULT 14  
AX934436  
LOCUS  
DEFINITION Sequence 13 from Patent WO02074942.  
ACCESSION AX934436  
VERSION AX934436.1 GI:40641682  
KEYWORDS  
SOURCE  
ORGANISM  
Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE  
1  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
CDS  
Gilbert, M. and Wakarchuk, W.W.  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
Patent: WO 02074942-A 13 26-SEP-2002;  
National Research Council of Canada (CA)  
Location/Qualifiers  
1. .873  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1. .>873

/note="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
Jejuni O:36"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04171.1"  
/db\_xref="GI:40641683"  
translation="MKKVIAGNPSLKEIDYSLRPNDPVRNQYFEDKYLGGK  
CKTVYTPNPFQEQYTKHLIQNQEYETELIMCSNYQAHLENENFVKTFDYFPDA  
HUGDYFFKQLEFNAYFRHEIYFNQRTSGVTMCVAIALGKEIYLSGIDFYQNGS  
SYAFDTQENLEKLAPDFKNDRSHYGHKNTDIKALELEKTKYKLYCLCPNSLLA  
NFIELAPLNSNFI IQEKNYTKDILIPSEAYGKFSKNFKIKIKENVYVKLIKD  
LLRLPSDIKHVFKG"

ORIGIN

Query Match 99.3%; Score 869.8; DB 2; Length 873;  
Best Local Similarity 99.8%; Pred. No. 5.1e-229;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATCTGGAATGACCAAGTTTAAAGAAATTCATTATTCACAG 60  
DB 1 ATGAAAAAGTTATTATCTGGAATGACCAAGTTTAAAGAAATTCATTATTCACAG 60  
QY 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAATGCAAGACAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAATGCAAGACAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAGAAATCAAGAAATGAGACCGAACTAATATGTTCTAAT 240  
DB 181 ACTTTAAACATTTAATCAGAAATCAAGAAATGAGACCGAACTAATATGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTTATTTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTTATTTTCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360  
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTAATGAGGCTCTATATGTCAGTAGCC 420  
DB 361 TTTACGAAATTTATTTCAATCAAGAAATTAATGAGGCTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAACTTACAAATAAATCTATATGCTTATGTCCTAATAGTCTTTTACCAATTTT 660  
DB 601 GAAAACTTACAAATAAATCTATATGCTTATGTCCTAATAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720  
DB 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720  
QY 721 AAAGATATATCATACCTTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780  
DB 721 AAAGATATATCATACCTTTCTAGTAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780  
QY 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATATAAAGATCTATTAAAGATTA 840  
DB 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATATAAAGATCTATTAAAGATTA 840

DB 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGCATATAAGCATTTATTTCAAGGAAAA 873  
DB 841 CCTAGCATATAAGCATTTATTTCAAGGAAAA 873  
RESULT 15  
AR481787 AR481787 873 bp DNA linear PAT 14-MAY-2004  
LOCUS Sequence 13 from patent US 6699705.  
DEFINITION AR481787  
ACCESSION AR481787  
VERSION AR481787.1 GI:47243422  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclonable.  
REFERENCE 1 (bases 1 to 873)  
AUTHORS Gilbert M. and Wakarchuk W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
JOURNAL Patent: US 6699705-A 13 02-MAR-2004;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES Location/Qualifiers  
source 1..873  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 99.3%; Score 869.8; DB 2; Length 873;  
Best Local Similarity 99.8%; Pred. No. 5.1e-229;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATCTGGAATGACCAAGTTTAAAGAAATTCATTATTCACAGG 60  
DB 1 ATGAAAAAGTTATTATCTGGAATGACCAAGTTTAAAGAAATTCATTATTCACAGG 60  
QY 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAATGCAAGACAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAATGCAAGACAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAGAAATCAAGAAATGAGACCGAACTAATATGTTCTAAT 240  
DB 181 ACTTTAAACATTTAATCAGAAATCAAGAAATGAGACCGAACTAATATGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTTATTTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTGTAAAACTTTTACGATTTATTTTCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360  
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTAATGAGGCTCTATATGTCAGTAGCC 420  
DB 361 TTTACGAAATTTATTTCAATCAAGAAATTAATGAGGCTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy	601	GAAGAACTTACAAAAATAAAACTATATTTGCTTATGTCTTAATAGTCTTTTATAGCAAAATTTT	660
Db	601	GAAGAACTTACAAAAATAAAACTATATTTGCTTATGTCTTAATAGTCTTTTATAGCAAAATTTT	660
Qy	661	ATAGAACTAGCGCCAAATTTAAATTCCAAATTTTATCATACAAGAAAGAAAAATAAAGTACACT	720
Db	661	ATAGAACTAGCGCCAAATTTAAATTCCAAATTTTATCATACAAGAAAGAAAAATAAAGTACACT	720
Qy	721	AAAGATATACTCATACCTCTCTAGTGGAGCTTATGGAAAAATTTTCCAAAAAATATTAATTTT	780
Db	721	AAAGATATCTCATACCTCTCTAGTGGAGCTTATGGAAAAATTTTCCAAAAAATATTAATTTT	780
Qy	781	AAAAAATAAAAAATAAAGAAATGTTTTTATCAAGTTGATAAAGATCTATTAAGATTAA	840
Db	781	AAAAAATAAAAAATAAAGAAATGTTTTTATCAAGTTGATAAAGATCTATTAAGATTAA	840
Qy	841	CTTAGTCATATAAAGCAATTTATTTCAAAGGAAAA	873
Db	841	CTTAGTCATATAAAGCAATTTATTTCAAAGGAAAA	873
RESULT	16		
AR527386			
LOCUS	AR527386	873 bp	DNA linear PAT 08-OCT-2004
DEFINITION	Sequence 13 from patent US 6723545.		
ACCESSION	AR527386		
VERSION	AR527386.1		
KEYWORDS	GI:53914407		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 873)		
TITLE	Gilbert,M. and Wakarchuk,W.W.		
JOURNAL	Polypeptides having .beta.-1,4-GalNac transferase activity		
	Patent: US 6723545-A 13 20-APR-2004;		
	National Research Council of Canada; Ottawa;		
	CAX;		
FEATURES	Location/Qualifiers		
source	1..873		
ORIGIN	/organism="unknown"		
	/mol_type="genomic DNA"		
Query Match	99.3%;	Score 869.8;	DB 2; Length 873;
Best Local Similarity	99.8%;	Pred. No. 5.1e-229;	
Matches	871;	Conservative 0;	Mismatches 2; Indels 0; Gaps 0;
Qy	1	ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTGATTATTC	60
Db	1	ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTGATTATTC	60
Qy	61	CTACCAATGATTTTGATGTATTTAGATGTAAATCAATTTTATTTTGAAGATAAATCTAT	120
Db	61	CTACCAATGATTTTGATGTATTTAGATGTAAATCAATTTTATTTTGAAGATAAATCTAT	120
Qy	121	CTTGGTAAAAATGCAAAAGCAGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC	180
Db	121	CTTGGTAAAAATGCAAAACAGTGTTTTCACACCCCTAAATTTCTCTTTGAGCAATACTAC	180
Qy	181	ACTTTAAAAACATTTTAATCCAAAAATCAAGAATATCAGACCGAACTAAATTTATGTGTTCTAAT	240
Db	181	ACTTTAAAAACATTTTAATCCAAAAATCAAGAATATCAGACCGAACTAAATTTATGTGTTCTAAT	240
Qy	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCCCT	300
Db	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTAATTTTCCCT	300
Qy	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360
Db	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360
Qy	361	TTTTCAGAAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGGCC	420

QY	181	AC	TTTAAAA	CAT	TTTAA	TCCAAAA	TATCAGAA	TATGAGAC	CGAAT	TATGTTG	TCTTAAT	240
DB	181	AC	TTTAAAA	CAT	TTTAA	TCCAAAA	TATCAGAA	TATGAGAC	CGAAT	TATGTTG	TCTTAAT	240
QY	241	TACA	CCAGCT	CAT	CTAGAAA	TGAAAAT	TTTCTG	TAAAC	CTTTTAC	GATAT	TTTTCCT	300
DB	241	TACA	CCAGCT	CAT	CTAGAAA	TGAAAAT	TTTCTG	TAAAC	CTTTTAC	GATAT	TTTTCCT	300
QY	301	GATG	CTCAT	TTGGGAT	TGAT	TTTTTTTAA	CAAC	TTAAAGAA	TTTAA	TGCTTAT	TTTAAA	360
DB	301	GATG	CTCAT	TTGGGAT	TGAT	TTTTTTTAA	CAAC	TTAAAGAA	TTTAA	TGCTTAT	TTTAAA	360
QY	361	TTT	CAGAAA	TTTAT	TTTCAAT	CAAGAA	TTTAC	CTCAGGG	CTATAT	GTGTG	CAAGTAC	420
DB	361	TTT	CAGAAA	TTTAT	TTTCAAT	CAAGAA	TTTAC	CTCAGGG	CTATAT	GTGTG	CAAGTAC	420
QY	421	ATAG	CCCTAG	GATACA	AGAAAT	TTTCTT	TCGGGA	TTTGAT	TTTTTAT	CAAAA	TGGGTCA	480
DB	421	ATAG	CCCTAG	GATACA	AGAAAT	TTTCTT	TCGGGA	TTTGAT	TTTTTAT	CAAAA	TGGGTCA	480
QY	481	TC	TTATG	CTTTT	GCATAC	CAAA	CAAGAA	ATCTTTT	TAAAC	TAGCC	CTGAT	540
DB	481	TC	TTATG	CTTTT	GCATAC	CAAA	CAAGAA	ATCTTTT	TAAAC	TAGCC	CTGAT	540
QY	541	GATG	CTCG	CACTAT	TATCGG	CACTAG	TAAAA	TACAGAT	TAAAG	CTTTAGAA	TTTCTA	600
DB	541	GATG	CTCG	CACTAT	TATCGG	CACTAG	TAAAA	TACAGAT	TAAAG	CTTTAGAA	TTTCTA	600
QY	601	GA	AAAACT	TACAAA	TAAACT	TATTTG	CTTATG	CTCTA	TATG	CTTTT	TAGCAAA	660
DB	601	GA	AAAACT	TACAAA	TAAACT	TATTTG	CTTATG	CTCTA	TATG	CTTTT	TAGCAAA	660
QY	661	ATAGA	CTAG	CGCCAA	ATTTAA	ATTTCA	ATTTTAT	CAFA	CAAGAAAA	AAAA	TAACTAC	720
DB	661	ATAGA	CTAG	CGCCAA	ATTTAA	ATTTCA	ATTTTAT	CAFA	CAAGAAAA	AAAA	TAACTAC	720
QY	721	AA	GATAT	ACTCAT	ACCTTCTAG	TGAGG	CTTATG	GAAAA	TTTTT	CAAAA	ATA	780
DB	721	AA	GATAT	ACTCAT	ACCTTCTAG	TGAGG	CTTATG	GAAAA	TTTTT	CAAAA	ATA	780
QY	781	AAAA	AAAA	TAAAA	TAAAG	AAAA	TGTTTAT	TATCA	AGTTG	TGATA	AAAGAT	840
DB	781	AAAA	AAAA	TAAAA	TAAAG	AAAA	TGTTTAT	TATCA	AGTTG	TGATA	AAAGAT	840
QY	841	CCTAG	TGAT	ATA	AAAG	CAAT	TTATTT	CAAA	AGGAAA	873		
DB	841	CCTAG	TGAT	ATA	AAAG	CAAT	TTATTT	CAAA	AGGAAA	873		

RESULT 18  
 AR689943 LOCUS  
 AR689943 Sequence 13 from patent US 6905867.  
 AR689943 ACCESSION  
 AR689943.1 GI:74471951  
 AR689943.1  
 .  
 UNKNOWN. SOURCE  
 UNKNOWN.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 873)  
 AUTHORS Gilbert,M. and Wakarchuk,W.W.  
 TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl  
 transferase activity  
 JOURNAL Patent: US 6905867-A 13 14-JUN-2005;  
 National Research Council of Canada; Ottawa;  
 CAX;  
 FEATURES Location/Qualifiers  
 1..873  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN

```
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 873)
TITLE Gilbert,M. and Wakarchuk,W.W.
JOURNAL Nucleic acids encoding beta.-1,4-GalNAC transferase
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
ORIGIN
1..873
/mol_type="genomic DNA"

Query Match 99.3%; Score 869.8; DB 2; Length 873;
Best Local Similarity 99.8%; Pred. No. 5.1e-229;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGMAAAAGTTATTATTCGTGAAATGACCAAGTTTAAAGAAATTTGATTATTCGAAGG 60
Db 1 ATGMAAAAGTTATTATTCGTGAAATGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60

QY 61 CTACCAATGATTTTGTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Db 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATCAATATGATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATCAATATGATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCTCT 300
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCTCT 300

QY 301 GATGCTCAITTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
Db 301 GATGCTCAITTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360

QY 361 TTTACGAAATTTATTTTCAATCAAGAAATACCTCAGGGTCTATATGCTGAGTAGCC 420
Db 361 TTTACGAAATTTATTTTCAATCAAGAAATACCTCAGGGTCTATATGCTGAGTAGCC 420

QY 421 ATAGCCTTAGGATACAAAGAAATTTATCTTTGGGAATTCATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCTTAGGATACAAAGAAATTTATCTTTGGGAATTCATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Db 481 TCTTATGCTTTTGTATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540

QY 541 GATCGCTCGACTATATCGCATAGTAAATACAGATATAAAGCTTTAGATTTCTTA 600
Db 541 GATCGCTCACATATATCGCATAGTAAATACAGATATAAAGCTTTAGATTTCTTA 600

QY 601 GAAAACTTACAAAATAAATCTATATGCTTATGCTTAAATAGTCTTTTACGCAATTTT 660
Db 601 GAAAACTTACAAAATAAATCTATATGCTTATGCTTAAATAGTCTTTTACGCAATTTT 660

QY 661 ATAGAATAAGCGCAATTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720
Db 661 ATAGAATAAGCGCAATTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720

QY 721 AAGATATATCATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATTAATTTT 780
Db 721 AAGATATATCATACCTCTAGTAGGCTTATGGAATTTTCAAAAATATTAATTTT 780

QY 781 AAAAAATAAATAAAGAAATGTTTATTAAGTGTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAATAAAGAAATGTTTATTAAGTGTGATAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAAGCATTATTTCACAGGAAAA 873
Db 841 CCTAGTGATATAAAGCATTATTTCACAGGAAAA 873

RESULT 20
AF215659 11474 bp DNA linear BCT 02-JAN-2001
LOCUS Campylobacter jejuni lipooligosaccharide biosynthesis locus,
DEFINITION partial sequence.
ACCESSION AF215659
VERSION AF215659.1 GI:12004275
KEYWORDS
SOURCE
ORGANISM Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.
TITLE The LOS biosynthesis locus of Campylobacter jejuni O:4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 11474)
AUTHORS Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1999) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A
0R6, Canada

FEATURES
source Location/Qualifiers
1..11474
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43432"
/serotype="O:4"
/db_xref="ATCC:43432"
/db_xref="taxon:197"
<1..360
/note="ORF-01; similar to RfaC from Helicobacter pylori"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAG43973.1"
/db_xref="GI:12004276"
/transl_table="MKNSDRIYLSLYILKFEVTPMPECILHFLALIVARITPYLNKK
HKKIIDNLQICFPQYTKERDKLSLIYENFAOFGIDCLONQNTTKKILNKVFIN
ENLLIDALAKRPIIFTTAHGNWEILSLAYAAKYCAISVIGKLSKSEMYELSONR
TQFDIELIDKGGIRQMLSAUKRAUGLITDQDCVERKSVRLKFNKRVNVMGASL
IAQRNALIIPVAYKEGKFCIEFFKAKDSQNSLELTLYOASCEEMIKKRPWEY
FFHRRFASVNEEYKGA"
1234..2490
/note="ORF-02; similar to WaaM from Helicobacter pylori"
/codon_start=1
/transl_table=11
/product="acyltransferase"
/protein_id="AAG43974.1"
/db_xref="GI:12004277"
/transl_table="MKNSDRIYLSLYILKFEVTPMPECILHFLALIVARITPYLNKK
HKKIIDNLQICFPQYTKERDKLSLIYENFAOFGIDCLONQNTTKKILNKVFIN
ENLLIDALAKRPIIFTTAHGNWEILSLAYAAKYCAISVIGKLSKSEMYELSONR
TQFDIELIDKGGIRQMLSAUKRAUGLITDQDCVERKSVRLKFNKRVNVMGASL
IAQRNALIIPVAYKEGKFCIEFFKAKDSQNSLELTLYOASCEEMIKKRPWEY
FFHRRFASVNEEYKGA"
1234..2490
/note="ORF-03; similar to LgtF from Neisseria
meningitidis"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAG43975.1"
/db_xref="GI:12004278"
/transl_table="MNLKQISVILVKNABQTLLECLNSLKDPEIILLNRSIDNTL
NIANKPKDFANLYIYHNAFTGFGALKNLALSYAKNDWILSDADEVLENECIEKLN
KIQEDNIIALSRLNLYKGEWIKACGMMWPDYVLRIFNKNFTFRNKNLHVESLVLPSNA
KKIYLKGLKHYSYKDIKSHLIDKQYYSLLAKQNKHKKSGVLKANLRAFVTFPRNYP
LKNGFLYKYGKFIISVCSALCTFFKMKLYELOROKPKTCALIITYNOKERLKLVD
SVNRLALFPLNEVLADGSKEDTARLIEYOKDFPCPLKHIWQDEGFKLSKSRNKT1
```



```
CDS
KNADSEYIIVIDGDMILEKDFIHEHLEFAQRKLFQGGSRVILNKKSESEILNKDDYRI
1FNKKDKFKSSKNFLAKIFYLSKKR"
2786..3955
/Note="ORF-04; local similarity (N-terminus) to Cps14J
from Streptococcus pneumoniae"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAG43976.1"
/db_xref="GI:12004279"
/translation="MKKIGWIPYVNVKYLRECLDSVINQTYTNLEIILVNDGSDPE
HSLNIAKTYTKDKRITTFDKNGGLSSARNIGIYFSGEYKLNKTKIHKENSLEIF
OLDGNNPYIKAYKVSQAQFNNKDLNFTYPSIDIYIPLDSNDYWKLNCEBCEVIRM
KNVDLWPHDCTCYEDNLIKHKHTRMEIFDEKCEIITPKEYANRALSVGSDISFG
WNGMDNFLKGLKFINFINEDIHFGIILIFASANKIYVSOKLYLCRLRANSISN
HDKKITKANVSYFKDIYETFTGENAKAKNYLKAASRVITALKLEFFKQDNENALA
IKETFLPCYAKKALMIKKFKDPLNKLKEQLVLKPIQTKPLDIWKFWQKIKNI"
4025..5068
/Note="CgtA"
/Genes="5068"
4025..5068
/Genes="cgtA"
/Note="glycosyltransferase; similar to CgtA from
Campylobacter jejuni OH4384 encoded by GenBank Accession
Number AF130984"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylgalactosaminyltransferase"
/protein_id="AAG43977.1"
/db_xref="GI:12004280"
/translation="MLFQSYFVKIICLIPFRKIRHKIKKTFLLKNIQRIKDISVLKP
KTLQVKNYNNEDLIKANKAIIRREGHKGFVNDKSKPKSLNPAWIRVKNNEAITL
KASLSILPAIORGVIGNDCTDGEIILECKQYSPFIPKIPYELIQIKNPKESEN
KLVSYNVVASIPKDEMLIKIDVDHIYDAKKLYKSYFIPKKNIDVWSYRSDIHTFN
DNFFLCKDNNGLPEPGDLLNNYNLKWKEVLDIRNNWKKATKQSFSSNIHSL
QLKXKRIILFHTLELNNYHPFLKQRAQDIYKYNWISIEBFKKFVLQNIHNKIEPSMI
SKETLKIKPLTLF"
complement (5054..5959)
/Genes="cgtB"
complement (5054..5959)
/Genes="cgtB"
/Note="Glycosyltransferase; similar to CgtB from
Campylobacter jejuni OH4384 encoded by GenBank Accession
Number AF130984"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAG43978.1"
/db_xref="GI:12004281"
/translation="WPKISIIILPTYNVEQYIARATESCINOTFKDIEIIVDDCGNDN
SINIAKEYSKDKRIKIHNEKNLGLLRARYEGVKVANSFYIMFLDPDDIYELNACEE
CIIKILDEQEDLVFNAIVEGNSVSYKXKDFNSGYSKKEFVKKIIAKKNLYWNGW
KLIRKLYLEAFASLKEKDVKNMAEDVLLYPLMSQAQKIAYMNCNIYHYVPPNNS
ICNTNKEVIVKNNIIOBLQVLNLYRQNYILNKYCSVLYVLYIKYLIQYIKRTKLM
VTLAKINILTLKILFKYKKFLQKC"
6048..6923
/Genes="cst-II"
6048..6923
/Genes="cst-II"
/Note="sialyltransferase; activity determined
experimentally"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAG43979.1"
/db_xref="GI:12004282"
/translation="MKKVIIAGNGPSLKEIDYSRLPNDPDRCNQFYEDKYKLGKK
CKAVFYTFPFEEQYITLKHILQNBETELMCSYNQAHLENFVKTFDYDFDA
HLGYDFFKQLEKNAYFKPEHIFYNQRTISGVYMCVAVATGALGYEIKLYSGIDYQNGS
SYAFDTKQNLKLPADPKNDRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLA
NFIELAPNLNSNFIIOEKNNYTKDILIPSEAYGKPSKNINPKKIKIKENVYKLIKID
LLRLPSDIKHYFKGK"
6324..7964
/Note="ORF-08; similar to SiaC from Neisseria
```

```
meningitidis"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAG43980.1"
/db_xref="GI:12004283"
/translation="MKEIKIQNIIISEEKAPLVVPEIGINHSLELAKIMVDAAFSA
GAKIIKHOTHIVEDEMSAAKQKVIIPNAKISIVEIMOKCALNYDEKALKEYTEKGLG
VYLSTPFSRAGVNRLEDMGVSAFKIGSGECNNVPLIKHIAAFKPKPMIVSTGMNSTESI
KPTVKILLDNEIPFVLMHTTNLYPTPHNLVRLNAMLLELKKFSCWGLSDHTONLAC
LGVALGALCERHRTDMSRSGDPIVCSMDTQALKELIIQSEQMAWRGNESKAA
KQSQVITDFAFASWSIKDIKKGEVLSMDNIIWKRPGLGGSASEFENILKKALRDI
ENDTQISYEDFT"
8021..9079
/Note="ORF-09; similar to SiaA from Neisseria
meningitidis"
/codon_start=1
/transl_table=11
/product="sialic acid biosynthesis protein"
/protein_id="AAG43981.1"
/db_xref="GI:12004284"
/translation="MYRVQNSSEFELIFATGMHLSKNFGYTVKELYKNGKNIYEPI
NYDKYQTDKALATTIDGFSRYVNELKPDILIVHGDRIEPLAAAVIGALNNILVAHIE
GGEISGTIDDSLRHAISKLAHILHWNDEFARLMQLEDKESFIIIGSPLELNDN
KISLSBAKYVDINENYALIMHPVTTEITSIKQADNLVKALISKNKIYIIVPNN
DILGSELLOSVEELKNNPRFKLPSLRFEYFITLLKNADFIIGNSSCTLKEALYLKTA
GILVGSQRQGRGLNGENTLKVANSDEILKALINTHIKQDLFSAKLEILDSSKLPFEYL
LSGEFFKLNQTKVFKDIK"
9076..9741
/Note="ORF-10; similar to NeuA from Haemophilus ducreyi"
/codon_start=1
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AAG43982.1"
/db_xref="GI:12004285"
/translation="MSLAIIPARGSGKIRKNLVLNNKPLIYTIKAALNTKSISK
VVVSSDSEILNVAKSNVDILKRPISLAQDITTSKVLHLALFKYQEDVVFLOPT
SPLRNTHIDEAFNLKYSNNALISVSECDNKILKAFVCDYDGLAGICNDYEPFIP
RQKLPTYMNGAIYIILKEFLNPNPFLQNKTHFLMDESSSLDIDCLDLKQVEQI
WKK"
9729..10562
/Note="ORF-11; similar to putative ORF from Rhodobacter
capsulatus encoded by GenBank Accession Number AF010496"
/codon_start=1
/transl_table=11
/product="acetyltransferase"
/protein_id="AAG43983.1"
/db_xref="GI:12004286"
/translation="MEKITLCKNKLNLKQYNIYTKYTIENPRFRSLTKTKDFTIT
PLENNQLESVAGIGIEYAKFSNIIHEMGSFSGSFLPHYAKVGRVCSYADGVSM
FNOQHPMDRISTASTFYETNHSFINNACONHINKTFPIVHNPNSSITHLIIQDDVMI
GKDVLLKQGITLGTGCVIGQRAVTKDVPVAVAGIPAKIIKYRKNFKNTIERLUKIQ
WVWYHFAFDYDIDILNLKINQYLDLLEKIIKKSIYNNPNKLYFRDILDELKSKIFNL
F"
complement (10554..11366)
/Note="ORF-12; similar to putative ORF HI0868 from
Haemophilus influenzae encoded by GenBank Accession Number
U32768"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAG43984.1"
Query Match 99.3%; Score 869.6; DB 15; Length 11474;
Best Local Similarity 99.5%; Pred. No. 1.9e-229;
Matches 872; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db 6048 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 6107
Qy 61 CTACCAATGATTTGATGATTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
|||||
```

Db	6108	CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	6167
Qy	121	CTTGGTAAATAAGCAAGAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACCTAC	180
Db	6168	CTTGGTAAATAAGCAAGAGTGTTTTACACCCCTGGTTCCTTTCTTTGAGCAATACCTAC	6227
Qy	181	ACTTTAAACATTTAATCCAAATATCAAGAAATATGAGACCGAACTAATATGTTCTTAAT	240
Db	6228	ACTTTAAACATTTAATCCAAATATCAAGAAATATGAGACCGAACTAATATGTTCTTAAT	6287
Qy	241	TACAAACAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT	300
Db	6288	TACAAACAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT	6347
Qy	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAACAACTTAAACAACTTAAACAACT	360
Db	6348	GATGCTCATTTGGGATATGATTTTAAACAACTTAAACAACTTAAACAACTTAAACAACT	6407
Qy	361	TTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC	420
Db	6408	TTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC	6467
Qy	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAAATGGGTCA	480
Db	6468	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAAATGGGTCA	6527
Qy	481	TCATTATGCTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Db	6528	TCATTATGCTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	6587
Qy	541	GATGCTCGACATATCGGACATATGTAATAACATGTAATAACATGTAATAACATGTAATA	600
Db	6588	GATGCTCGACATATCGGACATATGTAATAACATGTAATAACATGTAATAACATGTAATA	6647
Qy	601	GAATAAATCTTACAAAATTAATAATTTATGCTTATGCTTATGCTTATGCTTATGCTTATG	660
Db	6648	GAATAAATCTTACAAAATTAATAATTTATGCTTATGCTTATGCTTATGCTTATGCTTATG	6707
Qy	661	ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAAAGAAATAAATACTACACT	720
Db	6708	ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAAAGAAATAAATACTACACT	6767
Qy	721	AAAGATATCTATCTATCTCTAGGCTTATGGAATTTTCAAAAATTAATAATTTT	780
Db	6768	AAAGATATCTATCTATCTCTAGGCTTATGGAATTTTCAAAAATTAATAATTTT	6827
Qy	781	AAAAAATAAATAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA	840
Db	6828	AAAAAATAAATAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA	6887
Qy	841	CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 876	
Db	6888	CCTAGTGATATAAGCATTTATTTCAAGGAAATAA 6923	

RESULT 21  
AX934434  
LOCUS AX934434  
DEFINITION Sequence 11 from Patent WO02074942.  
ACCESSION AX934434  
VERSION AX934434.1 GI:40641680  
KEYWORDS  
SOURCE  
ORGANISM  
Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
JOURNAL Patent: WO 02074942-A 11 26-SEP-2002;  
National Research Council of Canada (CA)  
Location/Qualifiers

source	1. .873 /organism="Campylobacter jejuni" /mol_type="unassigned DNA" /db_xref="taxon:197" 1. .>873 /note="unnamed protein product; Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C. jejuni O:4" /codon_start=1 /transl_table=11 /protein_id="CAF04170.1" /db_xref="GI:40641681" /translation="MKKVIAGNGPSLKEIDYSLRPNDFVRCNOFVEDKYVLGKK HGVYDFPKQFENYAFKHEIYFNQRTSGVMCAVAIALGYKEIYLSGIDFYQDGA SYAPDTQENLLKLAPDFKNDRSHVIGHSKNTDIKALEFLEKTKYIKLYCLCPNSLLA NFLEAPLNSNFILQEKNNYTKDILIPSEAYGKFSKNINFKKIKENYVYVKLIKD LLRLPSDIKHVFKGK"										Query Match Best Local Similarity 98.9%; Score 866.6; DB 2; Length 873; Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
CDS	QY	1	ATGAAAAAGTTATTTGCTGGAATGACCAAGCTTTAAAGAAATTGATTATTCAAGG	60							
	DB	1	ATGAAAAAGTTATTTGCTGGAATGACCAAGCTTTAAAGAAATTGATTATTCAAGG	60							
	QY	61	CTACCAAAATGATTTTCATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120							
	DB	61	CTACCAAAATGATTTTCATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120							
	QY	121	CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTGAGCAATACTATC	180							
	DB	121	CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTGTTTCTTTGAGCAATACTATC	180							
	QY	181	ACTTTAAAAATTTAAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTTAAT	240							
	DB	181	ACTTTAAAAATTTAAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTTAAT	240							
	QY	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTCCCT	300							
	DB	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTCCCT	300							
	QY	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA	360							
	DB	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA	360							
	QY	361	TTTTCAGGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTGTCAGATAGCC	420							
	DB	361	TTTTCAGGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTGTCAGATAGCC	420							
	QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTTTATCAAAAATGGGTCA	480							
	DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTTTATCAAAAATGGGTCA	480							
	QY	481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTTAAAACTAGCCCTGATTTTAAAAAT	540							
	DB	481	TCATTATGCTTTTGATACCAACCAAGAAATCTTTTTAAAACTAGCCCTGATTTTAAAAAT	540							
	QY	541	GATGCTCGACATATTCGGACATATGTAATAACATGTAATAACATGTAATAACATGTAATA	600							
	DB	541	GATGCTCGACATATTCGGACATATGTAATAACATGTAATAACATGTAATAACATGTAATA	600							
	QY	601	GAAAAAATCTTACAAAATTAATAATTTATGCTTATGCTTATGCTTATGCTTATGCTTATG	660							
	DB	601	GAAAAAATCTTACAAAATTAATAATTTATGCTTATGCTTATGCTTATGCTTATGCTTATG	660							
	QY	661	ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAAAGAAATAAATACTACACT	720							
	DB	661	ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAAAGAAATAAATACTACACT	720							
	QY	721	AAAGATATCTATCTCTAGTGGGCTTATGGAAAATTTTTCAAAAAATAATTAATTTT	780							

```
|||||
Db 721 AAAGATATACATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAATTAATAATTAAGAAAAATGTTTATTACAAGTTGATTAAGAAATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAAAATGTTTATTACAAGTTGATTAAGAAATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTAATTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTAATTTCAAGGAAAA 873

RESULT 22
AR481786
LOCUS AR481786 873 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 11 from patent US 6699705.
ACCESSION AR481786
VERSION AR481786.1 GI:47243421
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 11 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..873
/mol_type="genomic DNA"

ORIGIN
Query Match 98.9%; Score 866.6; DB 2; Length 873;
Best Local Similarity 99.5%; Pred. No. 3.9e-228;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGCGAAATGACCAAGTTTAAAGAAATTAAGTAATTCGAAG 60
Db 1 ATGAAAAAGTTATTATTCTGCGAAATGACCAAGTTTAAAGAAATTAAGTAATTCGAAG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTCTTTGAGCAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTCTTTGAGCAATACTAT 120

QY 121 CTGTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTGTGTAAAAATGCAAGCAGTGTTTTACACCCCTGTTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAGAACTTTTACGATTAATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAGAACTTTTACGATTAATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTAATTTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTAATTTTAA 360

QY 361 TTTTCAGAAATTTATTTCATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTTCAGAAATTTATTTCATCAAGAAATTAACCTCAGGGGTCTATATGTTGCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
|||||
```

```
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTCAAAAATTAATTTGCTTATGTCCTTAATAGTCTTTTACCAATTTT 660
Db 601 GAAAAAATTTCAAAAATTAATTTGCTTATGTCCTTAATAGTCTTTTACCAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAACTACACT 720
QY 721 AAAGATATACATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATACATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTAATTTT 780
QY 781 AAAAAATTAATAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTAATTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTAATTTCAAGGAAAA 873

RESULT 23
AR527385
LOCUS AR527385 873 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 11 from patent US 6723545.
ACCESSION AR527385
VERSION AR527385.1 GI:53914406
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 873)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 11 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..873
/mol_type="genomic DNA"

ORIGIN
Query Match 98.9%; Score 866.6; DB 2; Length 873;
Best Local Similarity 99.5%; Pred. No. 3.9e-228;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAAGTTATTATTGCTGGAATATGACCAAGTTTAAAGAAATTAAGTAATTCGAAG 60
Db 1 ATGAAAAAAGTTATTATTGCTGGAATATGACCAAGTTTAAAGAAATTAAGTAATTCGAAG 60

QY 61 CTACCAATGATTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTGTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180
Db 121 CTGTGTAAAAATGCAAGCAGTGTTTTACACCCCTGTTTCTTTTGGAGCAATACTAC 180

QY 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTAATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTAATTTTCT 300
```

QY	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360
Db	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360
QY	361	TTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC	420
Db	361	TTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC	420
QY	421	ATAGCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA	480
Db	421	ATAGCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA	480
QY	481	TCATTGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Db	481	TCATTGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
QY	541	GATGCTCGCCTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA	600
Db	541	GATGCTCGCCTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA	600
QY	601	GAAGAACTTACAAATAAAATATATGCTTATGCTCCTAATAGTCTTTTAGCAATTTT	660
Db	601	GAAGAACTTACAAATAAAATATATGCTTATGCTCCTAATAGTCTTTTAGCAATTTT	660
QY	661	ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT	720
Db	661	ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT	720
QY	721	AAAGATATACCTACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTT	780
Db	721	AAAGATATACCTACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTT	780
QY	781	AAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA	840
Db	781	AAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA	840
QY	841	CCTAGTCATATAAGCAATTTTCAAGGAAAA	873
Db	841	CCTAGTCATATAAGCAATTTTCAAGGAAAA	873
RESULT 24			
LOCUS	AR609664	873 bp	DNA
DEFINITION	Sequence 11 from patent US 6825019.		linear
ACCESSION	AR609664		
VERSION	AR609664.1	GI:56664964	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 873)		
AUTHORS	Gilbert M. and Wakarchuk, W.W.		
TITLE	Polypeptides having beta -1,3-galactosyl transferase activity		
JOURNAL	Patent: US 6825019-A 11 30-NOV-2004; National Research Council of Canada; Ottawa; CAX;		
FEATURES			
source	Location/Qualifiers		
	1..873		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	98.9%;	Score 866.6;	DB 2; Length 873;
Best Local Similarity	99.5%;	Pred. No. 3.9e-228;	
Matches 869;	Conservative	0; Mismatches	4; Indels 0; Gaps 0;
QY	1	ATGAAAAAGTTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG	60
Db	1	ATGAAAAAGTTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG	60
QY	61	CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120

Db	61	CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120
QY	121	CTTGTAATAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC	180
Db	121	CTTGTAATAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC	180
QY	181	ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATATGTTCTTAAT	240
Db	181	ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATATGTTCTTAAT	240
QY	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACCGATATTTTCCCT	300
Db	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACCGATATTTTCCCT	300
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA	360
Db	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA	360
QY	361	TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC	420
Db	361	TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC	420
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA	480
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA	480
QY	481	TCATTGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
Db	481	TCATTGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540
QY	541	GATGCTCGCCTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA	600
Db	541	GATGCTCGCCTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA	600
QY	601	GAAGAACTTACAAATAAAATATATGCTTATGCTCCTAATAGTCTTTTAGCAATTTT	660
Db	601	GAAGAACTTACAAATAAAATATATGCTTATGCTCCTAATAGTCTTTTAGCAATTTT	660
QY	661	ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT	720
Db	661	ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT	720
QY	721	AAAGATATACCTACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTT	780
Db	721	AAAGATATACCTACCTCTAGTAGGCTTTATGGAAAAATTTTCAAAAAATTTTAAATTT	780
QY	781	AAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA	840
Db	781	AAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA	840
QY	841	CCTAGTCATATAAGCAATTTTCAAGGAAAA	873
Db	841	CCTAGTCATATAAGCAATTTTCAAGGAAAA	873
RESULT 25			
LOCUS	AR689942	873 bp	DNA
DEFINITION	Sequence 11 from patent US 6905867.		linear
ACCESSION	AR689942		
VERSION	AR689942.1	GI:74471950	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 873)		
AUTHORS	Gilbert, M. and Wakarchuk, W.W.		
TITLE	Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity		
JOURNAL	Patent: US 6905867-A 11 14-JUN-2005; National Research Council of Canada; Ottawa; CAX;		
FEATURES			
	Location/Qualifiers		



Db	721	AAAGATATCTATCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT	780
Qy	781	AAAAAATAAAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA	840
Db	781	AAAAAATAAAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA	840
Qy	841	CCTAGTGATATAAGCATTTATTTCAAGGAAAA	873
Db	841	CCTAGTGATATAAGCATTTATTTCAAGGAAAA	873
RESULT 27			
AY681242		1303 bp	DNA
LOCUS			linear BCT 06-MAY-2005
DEFINITION			Campylobacter jejuni isolate Contig_16b Cj81-010 and Cj81-011
ACCESSION			AY681242
VERSION			AY681242.1
KEYWORDS			genes, partial cds.
SOURCE			1 GI:57790356
ORGANISM			Campylobacter jejuni subsp. jejuni 81-176
REFERENCE			Campylobacter jejuni subsp. jejuni 81-176
AUTHORS			Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.
TITLE			1 (bases 1 to 1303)
JOURNAL			Poly, F., Threadgill, D. and Stintzi, A.
PUBMED			Genomic diversity in Campylobacter jejuni: identification of C. jejuni 81-176-specific genes
REFERENCE			J. Clin. Microbiol. 43 (5), 2330-2338 (2005)
AUTHORS			15872262
TITLE			2 (bases 1 to 1303)
JOURNAL			Poly, F. and Stintzi, A.
PUBMED			Direct Submission
REFERENCE			Submitted (08-JUL-2004) Veterinary Pathobiology, Oklahoma State University College of Veterinary Medicine, 250 McIlroy Hall, Stillwater, OK 74078, USA
FEATURES			Location/Qualifiers
source			1..1303
			/organism="Campylobacter jejuni subsp. jejuni 81-176"
			/mol_type="genomic DNA"
			/strain="81-176"
			/isolate="Contig_16b"
			/db_xref="taxon:354242"
			complement(<1..361)
			/note="similar to beta-1,3-galactosyltransferase in Campylobacter jejuni OH4384"
CDS			/codon_start=1
			/transl_table=11
			/product="Cj81-010"
			/protein_id="AAW56103.1"
			/db_xref="GI:57790357"
			/translation="MFKISILPTYNVEQYIARAIESCINOTFKDIEIIVVDDCGNDN SINIAKEYSKDKRIKIHNENKLGILLRAREYGVKVNSPYIMFLDPDDYLELNACEE CIKILDSODEDLVFFNA"
			450..21303
			/note="similar to alpha-2,3-sialyltransferase in Campylobacter jejuni OH4384"
CDS			/codon_start=1
			/transl_table=11
			/product="Cj81-011"
			/protein_id="AAW56104.1"
			/db_xref="GI:57790358"
			/translation="MKKVIIAGNGFSLKEIDYSRLPNDFVRCNQYFEDKYKLGKK KGVYFPNFFEGYTLKHLIQNEYETELIMCSNVOAHLENENFVKTFYFPDA HLGVDFYFKLFKFEIYFNQRTISGVYCAVAIALGYKEIYLSGIDFYONGS SYAFDTKOENLLKLPDKNDRSHVIGHSKNTDIKALEFLEKTKIKLYCICPNLSLLA NFIELAPNLSNFIQENKNTYKIDILIPSSRAYGFKSNINFKIKIKENYVYKLIKD LLRLPSDI"
			450..21303
ORIGIN			
Query Match			97.3%; Score 852.4; DB 15; Length 1303;
Best Local Similarity			99.9%; Pred. No. 2.7e-224;
Matches 853; Conservative			0; Mismatches 1; Indels 0; Gaps 0;
1			
Qy	1	ATGAAAAAGTTATTATTGCTCGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG	60
Db	450	ATGAAAAAGTTATTATTGCTCGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG	509
Qy	61	CTACCAATGATTTTGTAGTATTAATCAATTTTATTGAGAGATAAATCTAT	120
Db	510	CTACCAATGATTTTGTAGTATTAATCAATTTTATTGAGAGATAAATCTAT	569
Qy	121	CTTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATCTAC	180
Db	570	CTTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATCTAC	629
Qy	181	ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAATTAATATGTGTCTTAAT	240
Db	630	ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAATTAATATGTGTCTTAAT	689
Qy	241	TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCT	300
Db	690	TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCT	749
Qy	301	GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA	360
Db	750	GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA	809
Qy	361	TTTACAGAAATTTAATCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC	420
Db	810	TTTACAGAAATTTAATCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC	869
Qy	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
Db	870	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	929
Qy	481	TCTTATGCTTTGATACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT	540
Db	930	TCTTATGCTTTGATACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT	989
Qy	541	GATGCTCGCACTATATCGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA	600
Db	990	GATGCTCGCACTATATCGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA	1049
Qy	601	GAATAAACTTACAAATAAACTATATCTTATGCTTAAATAGTCTTTTAGCAAAATTTT	660
Db	1050	GAATAAACTTACAAATAAACTATATCTTATGCTTAAATAGTCTTTTAGCAAAATTTT	1109
Qy	661	ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATACTACT	720
Db	1110	ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATACTACT	1169
Qy	721	AAAGATATCTATCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTT	780
Db	1170	AAAGATATCTATCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTT	1229
Qy	781	AAAAAATAAAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA	840
Db	1230	AAAAAATAAAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA	1289
Qy	841	CCTAGTGATATAAA	854
Db	1290	CCTAGTGATATAAA	1303
RESULT 28			
AX934427			Sequence 4 from Patent WO02074942.
LOCUS			AX934427
DEFINITION			AX934427
ACCESSION			AX934427
VERSION			AX934427.1
KEYWORDS			GI:40641674
SOURCE			Campylobacter jejuni
ORGANISM			Campylobacter jejuni
			Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.
REFERENCE			1

Db 121 CTTGGTAAAAANTTCAAAGCAGTATTTACATCCTGGTCTTTTTTTTGAACAATACTAC 180



```
Qy 181 ACTTTAAACATTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTGTTCTAAT 240
Db |||||||
Qy 181 ACTTTAAACATTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTGTTCTAAT 240
Db |||||||
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTATAGGATATTTTCCT 300
Db |||||||
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTATAGGATATTTTCCT 300
Db |||||||
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db |||||||
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db |||||||
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTCAGTAGCC 420
Db |||||||
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGAGCTATATGTGTCAGTAGCT 420
Db |||||||
Qy 421 ATAGCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGGTCA 480
Db |||||||
Qy 421 ATAGCCTAGGATACAAAGAAATTTATCTTTCTGGAATGATTTTATCAAAATGGGTCA 480
Db |||||||
Qy 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db |||||||
Qy 481 TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db |||||||
Qy 541 GATCGCTCGCACTATATCGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Db |||||||
Qy 541 GATCGCTCACACTATATCGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Db |||||||
Qy 601 GAAAAAATTTACAAAAATTAAGAAATTTGCTTATGCTCCTAATAGTCTTTTAGCAAAATTT 660
Db |||||||
Qy 601 GAAAAAATTTACAAAAATTAAGAAATTTGCTTATGCTCCTAATAGTCTTTTAGCAAAATTT 660
Db |||||||
Qy 661 ATAGAACTAGGGCAATTTAAATTTAAATTTTATCATACAGAAAAATTAACCTACACT 720
Db |||||||
Qy 661 ATAGAACTAGGGCAATTTAAATTTAAATTTTATCATACAGAAAAATTAACCTACACT 720
Db |||||||
Qy 721 AAAGATATCTACACTCTCTAGCTGAGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 780
Db |||||||
Qy 721 AAAGATATCTACACTCTCTAGCTGAGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 780
Db |||||||
Qy 781 AAAAAATTAATAATTAAGAAATTTTATTTATCAAGTTGATAAAGATCTATTAAAGATTA 840
Db |||||||
Qy 841 CCTAGTATATAAGCAATTTTCAAGAGAAATTA 876
Db |||||||
Qy 841 CCTAGTATATAAGCAATTTTCAAGAGAAATTA 876
Db |||||||
```

## RESULT 30

```
AR271701
LOCUS AR271701 876 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 4 from patent US 6503744.
ACCESSION AR271701
VERSION AR271701.1 GI:29703246
KEYWORDS
SOURCE
ORGANISM Unknown.
```

## REFERENCE

```
1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6503744-A 4 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;
```

## FEATURES

```
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
```

## ORIGIN

Query Match 96.3%; Score 844; DB 2; Length 876;

```
Best Local Similarity 97.7%; Pred. No. 6.5e-222;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTATTCTCGAAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db |||||||
Qy 1 ATGAAAAAGTTATTATTCTCGAAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Db |||||||
Qy 61 CTACCAAAATGATTTTATAGTGTATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db |||||||
Qy 61 CTACCAAAATGATTTTATAGTGTATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db |||||||
Qy 121 CTTGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180
Db |||||||
Qy 121 CTTGTAAAAAATGCAAGCAGTGTTTTACCAATCTCTGCTTTTGTGAACAATACTAC 180
Db |||||||
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAATTAATGCTTCTAAT 240
Db |||||||
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAATTAATGCTTCTAAT 240
Db |||||||
Qy 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTTTCCT 300
Db |||||||
Qy 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTTTCCT 300
Db |||||||
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360
Db |||||||
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360
Db |||||||
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Db |||||||
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGAGCTATATGTGTCAGTAGCT 420
Db |||||||
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATGGGTCA 480
Db |||||||
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAATGATTTTATCAAAATGGGTCA 480
Db |||||||
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db |||||||
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
Db |||||||
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db |||||||
Qy 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db |||||||
Qy 601 GAAAAAATTTACAAAAATTAAGAAATTTGCTTATGCTCCTAATAGTCTTTTAGCAAAATTT 660
Db |||||||
Qy 601 GAAAAAATTTACAAAAATTAAGAAATTTGCTTATGCTCCTAATAGTCTTTTAGCAAAATTT 660
Db |||||||
Qy 661 ATAGAACTAGGGCAATTTTAAATTTAAATTTTATCATACAGAAAAATTAACCTACACT 720
Db |||||||
Qy 661 ATAGAACTAGGGCAATTTTAAATTTAAATTTTATCATACAGAAAAATTAACCTACACT 720
Db |||||||
Qy 721 AAAGATATCTACACTCTCTAGTGGGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 780
Db |||||||
Qy 721 AAAGATATCTACACTCTCTAGTGGGCTTATGCTTATGCTTATGCTTATGCTTATGCTT 780
Db |||||||
Qy 781 AAAAAATTAATAATTAAGAAATTTTATTTATCAAGTTGATAAAGATCTATTAAAGATTA 840
Db |||||||
Qy 781 AAAAAATTAATAATTAAGAAATTTTATTTTCAAGTTGATAAAGATCTATTAAAGATTA 840
Db |||||||
Qy 841 CCTAGTATATAAGCAATTTTCAAGAGAAATTA 876
Db |||||||
Qy 841 CCTAGTATATAAGCAATTTTCAAGAGAAATTA 876
Db |||||||
```

## RESULT 31

```
AR481783
LOCUS AR481783 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 4 from patent US 6699705.
ACCESSION AR481783
VERSION AR481783.1 GI:47243418
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
```



Unclassified.  
1 (bases 1 to 876)  
Gilbert, M. and Wakarchuk, W.W.  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
Patent: US 669705-A 4 02-MAR-2004;  
National Research Council of Canada; Ottawa;  
CAX;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

FEATURES  
source  
ORIGIN

Query Match 96.3%; Score 844; DB 2; Length 876;  
Best Local Similarity 97.7%; Pred. No. 6.5e-222;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGMAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60  
DB 1 ATGMAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGCTTTTGTGAACAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATTAATATGCTGCTTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATTAATATGCTGCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAAACTTTTAAAGAAATTTAATGCTTTATTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAAACTTTTAAAGAAATTTAATGCTTTATTTAAA 360

QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCT 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTTAAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840  
DB 601 GAAAAAATTTTAAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAGCAATTTATTTCAAAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAGCAATTTATTTCAAAGGAAAAATAA 876

RESULT 32  
AR527382 876 bp DNA linear PAT 08-OCT-2004  
LOCUS  
DEFINITION Sequence 4 from patent US 6723545.  
ACCESSION AR527382  
VERSION AR527382.1 GI:53914403  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity  
JOURNAL Patent: US 6723545-A 4 20-APR-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source  
ORIGIN

Query Match 96.3%; Score 844; DB 2; Length 876;  
Best Local Similarity 97.7%; Pred. No. 6.5e-222;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGMAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60  
DB 1 ATGMAAAAGTTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCGAAG 60

QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGCTTTTGTGAACAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATTAATATGCTGCTTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATTAATATGCTGCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAAACTTTTAAAGAAATTTAATGCTTTATTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAAACTTTTAAAGAAATTTAATGCTTTATTTAAA 360

QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCT 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTTAAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 660  
DB 601 GAAAAAATTTTAAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 660

```
|||||
601 GAAAAAATTACAAAATAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660
|||||
661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT 720
|||||
661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAGAAAAATAACTACACT 720
|||||
721 AAAGATATACATACACTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780
|||||
721 AAAGATATACATACACTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780
|||||
781 AAAAAATAAAAAATTTAAAGAAAAATTTTATACAAGTTGATAAAAGATCTATTAAAGATTA 840
|||||
781 AAAAAATAAAAAATTTAAAGAAAAATTTTATACAAGTTGATAAAAGATCTATTAAAGATTA 840
|||||
841 CCTAGTGATATAAAGCATTTTCAAGGAAAAATAA 876
|||||
841 CCTAGTGATATAAAGCATTTTCAAGGAAAAATAA 876
|||||

RESULT 33
AR609661
LOCUS AR609661 876 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 4 from patent US 6825019.
ACCESSION AR609661
VERSION AR609661.1 GI:56664961
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1. (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-l,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 4 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 96.3%; Score 844; DB 2; Length 876;
Best Local Similarity 97.7%; Pred. No. 6.5e-22;
Matches 856; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
QY 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAACTACTAT 120
DB 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAACTACTAT 120
QY 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
DB 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTTAAATCCAAAAATCAAGATATAGACCGAAGTAATTTATGTTCTAAT 240
DB 181 ACTTTAAACATTTTAAATCCAAAAATCAAGATATAGACCGAAGTAATTTATGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
QY 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTHAA 360
DB 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTHAA 360
QY 361 TTTTACGAAATTTTAAATCAAGATTAACCTCAGGGTCTATATGTTGCGAGTAGCC 420
DB 361 TTTTACGAAATTTTAAATCAAGATTAACCTCAGGGTCTATATGTTGCGAGTAGCT 420
```

```
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTCATTTTATCAAAATGGGTCA 480
|||||
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTCATTTTATCAAAATGGGTCA 480
|||||
QY 481 TCTTATGCTTTTGTATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
|||||
DB 481 TCTTATGCTTTTGTATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
|||||
QY 541 GATCCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
|||||
DB 541 GATCCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
|||||
QY 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
|||||
DB 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGTCCTAAATAGTCTTTTAGCAAAATTTT 660
|||||
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
|||||
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
|||||
QY 721 AAAGATATACATACACTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780
|||||
DB 721 AAAGATATACATACACTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780
|||||
QY 781 AAAAAATAAAAAATTTAAAGAAAAATTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
DB 781 AAAAAATAAAAAATTTAAAGAAAAATTTTATTAACAAGTTGATAAAAGATCTATTAAAGATTA 840
|||||
QY 841 CCTAGTGATATAAAGCATTTTCAAGGAAAAATAA 876
DB 841 CCTAGTGATATAAAGCATTTTCAAGGAAAAATAA 876
|||||

RESULT 34
AR689939
LOCUS AR689939 876 bp DNA linear PAT 12-SEP-2005
DEFINITION Sequence 4 from patent US 6905867.
ACCESSION AR689939
VERSION AR689939.1 GI:74471947
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1. (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity
JOURNAL Patent: US 6905867-A 4 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 96.3%; Score 844; DB 2; Length 876;
Best Local Similarity 97.7%; Pred. No. 6.5e-22;
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
QY 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAACTACTAT 120
DB 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAACTACTAT 120
QY 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
DB 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
```

```
QY 181 ACTTTAAACATTTAATCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
    |||
Db 181 ACTTTAAACATTTAATCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
    |||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
    |||
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
    |||
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAGAAATTAATGCTTATTTTAAA 360
    |||
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAGAAATTAATGCTTATTTTAAA 360
    |||
QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCT 420
    |||
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCT 420
    |||
QY 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
    |||
Db 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
    |||
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
    |||
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
    |||
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAATTTCTA 600
    |||
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAATTTCTA 600
    |||
QY 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660
    |||
Db 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660
    |||
QY 661 ATAGAACTAGCGCAATTTAAATTTCAAAATTTTATCATACAGAAATAAATACACACT 720
    |||
Db 661 ATAGAACTAGCGCAATTTAAATTTCAAAATTTTATCATACAGAAATAAATACACACT 720
    |||
QY 721 AAAGATATACCTACCTCTAGTGGCTTATGGAATTTTCAAAATAATTAATTTT 780
    |||
Db 721 AAAGATATACCTACCTCTAGTGGCTTATGGAATTTTCAAAATAATTAATTTT 780
    |||
QY 781 AAAAAATAAAATTTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
    |||
Db 781 AAAAAATAAAATTTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
    |||
QY 841 CCTAGTATAAAGCATTATTTCAAAGGAAATAA 876
    |||
Db 841 CCTAGTATAAAGCATTATTTCAAAGGAAATAA 876
    |||
```

```
RESULT 35
AR691835
LOCUS AR691835
DEFINITION Sequence 4 from patent US 6911337.
ACCESSION AR691835
VERSION AR691835.1 GI:74479546
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAc transferase
JOURNAL Patent: US 6911337-A 4 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
    source Location/Qualifiers
        1..876
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Query Match 96.3%; Score 844; DB 2; Length 876;
Best Local Similarity 97.7%; Pred. No. 6.5e-222;
```

```
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 1 ATGAAAAAGCTTATTTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
    |||
Db 1 ATGAAAAAGCTTATTTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
    |||
QY 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATFACTAT 120
    |||
Db 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATFACTAT 120
    |||
QY 121 CTTGGTAAAAATGCAAGAGAGTGTTTTACACCCCTAAATTTCTCTTTGAGCAATFACTAC 180
    |||
Db 121 CTTGGTAAAAATGCAAGAGAGTGTTTTACACCCCTAAATTTCTCTTTTGAACAATFACTAC 180
    |||
QY 181 ACTTTAAACATTTTAAATCCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
    |||
Db 181 ACTTTAAACATTTTAAATCCCAAAATCAAGAAATATGAGACCGAACTAAATATGTGTTCTAAT 240
    |||
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
    |||
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCCT 300
    |||
QY 301 GATGCTCATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
    |||
Db 301 GATGCTCATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
    |||
QY 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCT 420
    |||
Db 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGTCAGTAGCT 420
    |||
QY 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
    |||
Db 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
    |||
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
    |||
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
    |||
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAATTTCTA 600
    |||
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAGAATTTCTA 600
    |||
QY 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660
    |||
Db 601 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660
    |||
QY 661 ATAGAACTAGCGCAATTTAAATTTCAAAATTTTATCATACAGAAATAAATACACACT 720
    |||
Db 661 ATAGAACTAGCGCAATTTAAATTTCAAAATTTTATCATACAGAAATAAATACACACT 720
    |||
QY 721 AAAGATATACCTACCTCTAGTGGCTTATGGAATTTTCAAAATAATTAATTTT 780
    |||
Db 721 AAAGATATACCTACCTCTAGTGGCTTATGGAATTTTCAAAATAATTAATTTT 780
    |||
QY 781 AAAAAATAAAATTTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
    |||
Db 781 AAAAAATAAAATTTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
    |||
QY 841 CCTAGTATAAAGCATTATTTCAAAGGAAATAA 876
    |||
Db 841 CCTAGTATAAAGCATTATTTCAAAGGAAATAA 876
    |||
```

```
RESULT 36
AF400048
LOCUS AF400048
DEFINITION Campylobacter jejuni strain ATCC 43438 heptosyltransferase I (waaC)
    gene, partial cds; lipid A biosynthesis acyltransferase (htrB),
    putative two-domain glycosyltransferase, putative
    glycosyltransferase, beta-1,4-N-acetylglactosaminyltransferase
    (cgtA), beta-1,3-galactosyltransferase (cgtB),
    alpha-2,3-/alpha-2,8-sialyltransferase (cstII), putative sialic
    acid synthase (neuB), putative N-acetylglucosamine-6-phosphate
    11442 bp DNA linear BCT 03-JAN-2002
AF400048
LOCUS AF400048
DEFINITION
```

2-epimerase (neu1), CMP-Neu5Ac synthetase (neuA), putative acetyltransferase, and putative glycosyltransferase genes, complete cds; and heptosyltransferase II (waaf) gene, partial cds.

AF400048  
AF400048.1 GI:15193214

CDS

ACCESSION  
KEYWORDS  
SOURCE  
ORGANISM

Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

REFERENCE  
AUTHORS

1 (bases 1 to 11442)  
Gilbert M., Karwaski M. F., Bernatchez S., Young N. M., Taboada E.,  
Michniewicz J., Cunningham A. M. and Wakarchuk W. W.

TITLE

The Genetic Bases for the Variation in the Lipo-oligosaccharide of  
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF  
STALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE

JOURNAL  
PUBMED

J. Biol. Chem. 277 (1), 327-337 (2002)

REFERENCE  
AUTHORS

2 (bases 1 to 11442)  
Gilbert M., Michniewicz J., Karwaski M. F., Cunningham A. and  
Wakarchuk W. W.

TITLE  
JOURNAL

Submitted (15-JUL-2001) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada

FEATURES  
source

Location/Qualifiers  
1..11442  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="ATCC 43438"  
/db\_xref="ATCC:43438"  
/db\_xref="taxon:197"  
/notes="Fenner-type strain O:10"  
<1..348  
/genes="waaC"  
<1..348  
/genes="waaC"  
/codon\_start=1  
/transl\_table=11  
/product="heptosyltransferase I"  
/protein\_id="AAK91719.1"  
/db\_xref="GI:15193215"  
/translation="FAKEVLNLSGIDETHIELAPKFSLEELMAFTKMDLLIGNDSGP  
THUAFALNKASITIFGATPSYRNAFQTHINKIIDAGKQIONAKHIDKSDFCITRIEE  
DIFKLAKGLNEX"  
338..1225  
/genes="htrB"  
338..1225  
/genes="htrB"  
/codon\_start=1  
/transl\_table=11  
/product="lipid A biosynthesis acyltransferase"  
/protein\_id="AAK91720.1"  
/db\_xref="GI:15193216"  
/translation="MKNSDRIYLSLYILKFFVTFMEPECILHFLALIVARITFYLNKK  
HRKIIDINLQFPQYQKRDKLSKIYENAFQGDICLQNTTKKILKNKFNIN  
ENLLIDALKRPIIFTHAYHWEILSLAYAAKYGAISIVGKLSKSEIMVEILSQNR  
TQFDILIDKGGIGROMLSALKGRALGILTODCVKESVRLKFFNKENVYQMGASL  
IAORSNALITPVYAYKEDGKFCIEFFRVKDSQNASLEELTYQAQSCCEMIKKRPWEY  
FFHRFPASNYKIYDKAK"  
1222..2769  
/notes="orf3"  
/codon\_start=1  
/transl\_table=11  
/product="putative two-domain glycosyltransferase"  
/protein\_id="AAK91721.1"  
/db\_xref="GI:15193217"

gene

CDS

gene

CDS

complement (5043..5936)  
/gene="cgtB"  
complement (5043..5936)  
/gene="cgtB"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,3-galactosyltransferase"  
/protein\_id="AAK91724.1"  
/db\_xref="GI:15193220"

gene

CDS

translation="WFKSIIILPTVNEQVIARAIESCINQTFKNEHIIIVDDCGSDK  
SIDIVYKAKDDRIKIIHNEENLKLRLARYEGKVANSYIMFLPDDVLELNACSE  
CMKILKNEIDILLFPNAFVLENNKIERKINFOEKCYVKRDFKLLKTNLFWTWA  
KVIKKELYLKAAGLISLENAKINMAEDVLLYPLINISNTIFHLSKNLNYQINNF5I  
TKTLTLQNTKTNIOQSDNVLYLLKKMQYNNFNLTLLKLEIYFLLIEKYSLSKRNVL  
CFKINIFPKKIQKFYRLKKM"  
6025..6900  
/gene="cstII"  
6025..6900  
/gene="cstII"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-2,3-/alpha-2,8-sialyltransferase"  
/protein\_id="AAK91725.1"  
/db\_xref="GI:15193221"

gene

CDS

translation="MKNVILSGNPSLKEIDYSLRPNDPFRCNQFYPEDKYVYGKK  
PKAVPNPGLPFEQYTYTLKHLQNGEYETELIMCSNYNOAHLENEFVKTFDYFPDA  
HLGYDFKQLKEFNAYFPHEIYLLNQRITSGVTMCVAVALGKYEIYLSGIDFYQNGS  
SYAFDTQENLLKAPDFKNDKSHYIGHSEADYTDILKALEFLKTYKIKLYCLCPNSLLA  
NFIELAPNLSNFIIOEKNNYTKDILIPSEAYGKFSQINFPKKIKENIYYKLIKID  
LLRLPSDIKHYPKKGK"  
6901..7941  
/gene="neuB1"  
6901..7941  
/gene="neuB1"  
/codon\_start=1  
/transl\_table=11  
/product="putative sialic acid synthase"  
/protein\_id="AAK91726.1"

CDS

gene

CDS

translation="MKNVILSGNPSLKEIDYSLRPNDPFRCNQFYPEDKYVYGKK  
PKAVPNPGLPFEQYTYTLKHLQNGEYETELIMCSNYNOAHLENEFVKTFDYFPDA  
HLGYDFKQLKEFNAYFPHEIYLLNQRITSGVTMCVAVALGKYEIYLSGIDFYQNGS  
SYAFDTQENLLKAPDFKNDKSHYIGHSEADYTDILKALEFLKTYKIKLYCLCPNSLLA  
NFIELAPNLSNFIIOEKNNYTKDILIPSEAYGKFSQINFPKKIKENIYYKLIKID  
LLRLPSDIKHYPKKGK"  
6901..7941  
/gene="neuB1"  
6901..7941  
/gene="neuB1"  
/codon\_start=1  
/transl\_table=11  
/product="putative sialic acid synthase"  
/protein\_id="AAK91726.1"

KASKSEYIILIDGDMILEKDFVKNHLEFAKRVILQGSRVILNKNSEELKNNYSL  
AFNKKGNQRSIFIAKCVYFSKLTKKFKKSQLVKSGKTCNMSFYKSDFEAIEGN  
EKFTGWGREDSEFVARFLNDGVFKRLKFNALAYHYHEBNSKNMLEINHQIYLETIK  
NEKTWR"  
2773..3957  
/note="orf4"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAK91722.1"  
/db\_xref="GI:15193218"  
/translation="MKKVGWVPIYNVEKYLKECLDSVINQTYTNLEIILVNDGSTDE  
NSLNAKTYTLKDRITILFDKNGGLSSARNVIGIEFSGEYKLNKTQHTIKENSLIEF  
QLDGNPNYKAYKSOAFNNEKDLTNLTYSIDYIIIFLDSNYSKLNLCIEBCVIRM  
KNDVLFWDHDCITYEDNINKHKTRMEIFDFKKECIITPKEVANALSVGSDISFG  
WNGVIDFNLKQIKLKFNFILNEDIHFGIILPASANKIYVLSQKLYLCPLRANSIN  
HDKKITPCANSEYFKDIYETPGENAKEAKNYLKAASRYVLTALKLIEFFKQKSENALA  
IKETITPCYAKKALMIKKFKDPLNLKEQLVLIKPFIQTKLPYDIWKFQWOKIKYLLI  
KI"  
4013..5071  
/gene="cgtA"  
4013..5071  
/gene="cgtA"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,4-N-acetylgalactosaminyltransferase"  
/protein\_id="AAK91723.1"  
/db\_xref="GI:15193219"

gene

CDS

translation="MLFQSYFVKIICLPIPRKIRHKIKKTFLLKNQIRDKIDSYLPK  
KTLIQINKNEDLILKNAIIGGGHGFYNDKSDPKSPLNPAFIRVONEAITL  
KASLESILPAIQGVIGYNDCTDGBEIIIEFCQYPSPIPIKYVEIQIONPKSEEN  
KL5SYNYVASPIPKDEWLKIDVDHYDAKLYKSYIPRKNYHVISYRIDFIPNE  
EKFYVRNKEGELKAPGDCIAIQNTNLFWEKILIEDDTFKWNTAKNNIENAKSYEL  
KVRNRIYFTTELNNYHFFPIKNYKNDYKQLNWSLDDFIKNYKEKLNQIDFAMLEY  
KTLKKVYKKLTSASDCKI"  
complement (5043..5936)  
/gene="cgtB"  
complement (5043..5936)  
/gene="cgtB"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,3-galactosyltransferase"  
/protein\_id="AAK91724.1"  
/db\_xref="GI:15193220"



/gene="cst-II"  
/note="bifunctional sialyltransferase"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-2,3/8-sialyltransferase"  
/protein\_id="AAL36462.1"  
/db\_xref="GI:17221831"  
translation="MKKVI IAGNGPSLKEIDYSLRPNDFVRCNQYFEDKYVLGKK  
CKAYNPILPFEQVYTLKHLIQOEYTELIMCSYNOAHLENENFVKTVDFVFPDA  
HLYDFPQLKDFNAYKFEHLYFNQRIKTSYVMCAVAIALGYKEIYLSGIDFYQNGS  
SYADFTKQKLLKAPFNKDNDSHYIGHSKNDIKALEFLKTKYKIKYLCUCPSLLA  
NFIELAPLNSNFITIOEKNNYTKDILIPSSSEAYGKFSKNIINFKIKIKENIYYKLKID  
LLRLPSDIKHYFKG"

ORIGIN

Query Match 95.6%; Score 837.6; DB 15; Length 876;  
Best Local Similarity 97.3%; Pred. No. 3.8e-220;  
Matches 852; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60  
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTTCAAG 60  
QY 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180  
DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTCTTTTGAACAATACTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240  
QY 241 TACAAACAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300  
DB 241 TACAAACAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
QY 361 TTTACGAAATTTATTTCAATCAAGAATATACCTCAGGGGCTATATGTTGCAGTAGCC 420  
DB 361 TTTACGAAATTTATTTCAATCAAGAATATACCTCAGGGGCTATATGTTGCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTGCACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
DB 541 GATAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTCAAAATAAACTATTTGCTTATGCTCTAATAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTTCAAAATAAACTATTTGCTTATGCTCTAATAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGGCAAAATTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
DB 661 ATAGAACTAGCGGCAAAATTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
QY 721 AAAGATATCTATCATCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTATTAATTTT 780  
DB 721 AAAGATATCTATCATCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTATTAATTTT 780  
QY 781 AAAAAATAAAATTAAGAAATTTTATTTACAAATTTGATTAAGATCTATTAAAGATTA 840  
DB 781 AAAAAATAAAATTAAGAAATTTTATTTACAAATTTGATTAAGATCTATTAAAGATTA 840

Db 781 AAAAAATAAAATTAAGAAATTAATTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAATAA 876  
DB 841 CCTAGTGATATAAGCATTATTTCAAAGGAAATAA 876  
AY644679 6047 bp DNA linear BCT 01-JUL-2005  
Campylobacter jejuni strain CF 93-6 LOS outer core biosynthesis  
gene cluster, partial sequence.  
ACCESSION  
AY644679  
VERSION  
AY644679.1 GI:55420298  
KEYWORDS  
Campylobacter jejuni  
SOURCE  
ORGANISM  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE  
1 (bases 1 to 6047)  
Koga, M., Gilbert, M., Li, J., Koike, S., Takahashi, M., Furukawa, K.,  
Hirata, K. and Yuki, N.  
Antecedent infections in Fisher Syndrome: a common pathogenesis of  
molecular mimicry  
Unpublished  
2 (bases 1 to 6047)  
Gilbert, M., Koga, M., Karwaski, M.F., Leclerc, S. and Yuki, N.  
Direct Submission  
Submitted (02-JUN-2004) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada  
FEATURES  
Location/Qualifiers  
1..6047  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="CF 93-6"  
/serotype="HS:2"  
/db\_xref="taxon:197"  
57..1313  
/note="Orf1"  
/codon\_start=1  
/transl\_table=11  
/product="putative glucosyltransferase"  
/protein\_id="AAV52035.1"  
/db\_xref="GI:55420299"  
/translation="MNLKQISVIIIVKNAEQTLLLECLNSKDFEILLANNESIDNTL  
NIANKFKDFANLYIHNAFIFGALKNLALSYAKNDWILSIDADEVLENEICELKXN  
KLQSDNIALSRKLYKGWIKACGWPDYVLRIFNKPTRENDNLVHESLIVPSNA  
KKIYLNGLKHYSKYDIHSLIDKMOYSSLMKAKONTHKSGVLKANLRAFWTPFRNYF  
LKNGLYKGFIIISVCSALGTFFKPMKLYELQRPKTKCALLIITYNOKERLKLVD  
SVKNLAFLPNEVLIIADDSKEDTARLIEEYQKDFPCPLKHQWQDEBGFKLKSKRNKTL  
KNADSEYIIVIDGDMILEKDFIKHLEFAQRKLFLOGSRVLNKKSEELINKDDYRI  
IFNKDKFSKSNFLAKIFYSLSKKR"  
1609..2778  
/note="Orf2"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAV52036.1"  
/db\_xref="GI:55420300"  
/translation="MKKIGVPIYINVEKYLRECLDSVINOTYTNLEIILVNDGSTDE  
HSLNAKETLKDRTITLFDKNGGLSSARNIGIEYFSGEYKLNKQTHQKENSLEIF  
QLDGNPNYKAYKSSQFNNKQDITFTPSIDYIIFLSDNYWKLNCIEECVIRM  
KNVDWFDHDCITBEDNIKNHKKTMEIFDFKKECIITTPKEYNALSVGSRDISFG  
HGMKIDFNFLKOIKLFINFIINEDIHFGIILFASAKTYVLQKLYLCLRLANSIGN  
WDKKITKANVSEYKDIYETFGENAKAENYLKAAARVITATLKLIFEFKQKQENALA  
IKETPLPCYAKKALMIKKFKDPLNLKEQLVLKPFIOQLPYDINKFWQKIKNI"  
2848..3891  
/gene="cgtA"  
2848..3891  
/gene="cgtA"  
/codon\_start=1  
/transl\_table=11  
CDS  
gene  
CDS

```

/product="beta-1,4-N-acetylglucosaminyltransferase"
/protein_id="AAV52037.1"
/db_xref="GI:55420301"
/translation="MLFQSFVKIICLIFIPRKIRHKIKTFLILKNQORDKIDSLPK
KTLVQINKYNEDILKNAIIIEGHGKGYNYDEKSDPKSPKPNPWFPIRQNEAITL
KASLESILPAIRGVIGYNDCTGSESEIIEFCKQPSFPIKPYPIQIQNKSEEN
KLSYNYVASFIPKPEWLKIDNNHLYKWEKVIIDRINNKKATQKQSFSSNHSLE
DNFPLCKDNNNGNLIKPEQDCLLNNVNLKWEKVIDRINNKKATQKQSFSSNHSLE
OLYKXRIILPHTLNNYHFPFLKQRAQDIYKYNWISIEEFKFKYLNINHKIEPSMI
SKETLKKIFLTLF"
/gene="cgtB"
/complement(3877. .4782)
/gene="cgtB"
/complement(3877. .4782)
/gene="cgtB"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAV52038.1"
/db_xref="GI:55420302"
/translation="MKISILPTYNVEOYIARIESCINQTPKIDIEIIVVDDGNDN
SINIAKYSKKRIKIIHNEKNLGLLRAREYEGVKVANSYIMFLDPPDDLELNACEE
CIKILDEQSDVILVFNAINESVNIYSYKFDNSFGYSKFEFKVKKIIAKKNLYMWG
KLRLKYLFAFASLRLEKDKVINMAEDVLLIYPMLSQAQKIAYMNCNLHYHVPNNNS
ICNTKNEVLKNNIQLQVLNLYLRQNYILNKYCSVLVLYIKVLLYIYIKIKRTKLM
VTLAKNIIILTKILFKYKKFLQC"
/4871..5746
/gene="csc-II"
/4871..5746
/gene="cst-II"
/codon_start=1
/transl_table=11
/product="alpha-2,3-/2,8-sialyltransferase"
/protein_id="AAV52039.1"
/db_xref="GI:55420303"
/translation="MKVYIAGNPSLKEIDYSLRPNDPFCNPFYFEDKYYLGK
KCAVFNPIYLFQFYTLKHLIONOEYETELIMCSNNOAHLENEPVKTFYDFPDA
HGVDFPKOLKDNAYPKFHEIYNORITSGVMCAVAIALGVKEIYVLSGIDFVONGS
SYAFDQKQNLKLAENFNDNSHYIGHSKNTDIKALEFLEKTKIKLYLCIPNSLLA
NFIELAPNLSNFIYQKNNYTKDILIPSSSEAYGKFSKNINFKIKIKENIYYKLKD
LLRLPSDIKHYFKGK"
/5747..>6047
/gene="neuB"
/5747..>6047
/gene="neuB"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAV52040.1"
/db_xref="GI:55420304"
/translation="MKRIKIONIIIESEKAPLVVPEIGINHSLELAKIMVDAFST
GAKIIRHQTHIVEDMSKAARKVPIGNAKISIVIEIMQKCALDYKDELALKEYTEKL"

ORIGIN
Query Match          95.6%; Score 837.6; DB 15; Length 6047;
Best Local Similarity 97.3%; Pred. No. 1.7e-220; Indels 0; Gaps 0;
Matches 852; Conservative 0; Mismatches 24;

Qy 1 ATGAAAAAGTTATTGCTGGAATCGCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 4871 ATGAAAAAGTTATTGCTGGAATCGCAAGTTTAAAGAAATTTGATTATTCAAGA 4930

Qy 61 CTACCAATGATTTGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 4931 CTACCAATGATTTGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 4990

Qy 121 CTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTCTTGACATACTAC 180
Db 4991 CTGGTAAAAATGCAAGCAGTGTTTTACCAATCCTATTTCTTTTGTGAACAATACTAC 5050

Qy 181 ACTTTAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAAGTAATATGTTCTAAT 240
Db 5051 ACTTTAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAAGTAATATGTTCTAAT 5110

```

```

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTCT 300
Db 5111 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTCT 5170

Qy 301 GATGCTCATTTGGGATGATGATTTTAAACAACCTTAAAGAAATTTAAATCTTATTTTAA 360
Db 5171 GATGCTCATTTGGGATGATGATTTTAAACAACCTTAAAGAAATTTAAATCTTATTTTAA 5230

Qy 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
Db 5231 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 5290

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 5291 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 5350

Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAATAAT 540
Db 5351 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAATAAT 5410

Qy 541 GATCGCTCGACTATATCGGCATAGTAAATAACAGATATATAAGAGCTTTAGAAATTTCTA 600
Db 5411 GATAATTACACTATATCGGCATAGTAAATAACAGATATATAAGAGCTTTAGAAATTTCTA 5470

Qy 601 GAAAAAATCTTACAAAAATATAATTTGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660
Db 5471 GAAAAAATCTTACAAAAATATAATTTGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 5530

Qy 661 ATAGAATCTAGCGCCAAATTTAAATTTTCAATTTTATCATAACAAGAAAAATTAACACACT 720
Db 5531 ATAGAATCTAGCGCCAAATTTAAATTTTCAATTTTATCATAACAAGAAAAATTAACACACT 5590

Qy 721 AAGATATACATACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTATTTT 780
Db 5591 AAGATATACATACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTATTTT 5650

Qy 781 AAAAAATATAAAATTTAAAGAAATTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Db 5651 AAAAAATATAAAATTTAAAGAAATTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 5710

Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876
Db 5711 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 5746

RESULT 39
AX934424
LOCUS 11474 bp DNA linear PAT 05-JAN-2004
DEFINITION Sequence 1 from Patent WO02074942.
ACCESSION AX934424
VERSION AX934424.1 GI:40641671
KEYWORDS
SOURCE
ORGANISM
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1
AUTHORS
TITLE
JOURNAL
National Research Council of Canada (CA)
LOCATION/Qualifiers
1. .11474
/organism="Campylobacter jejuni"
/mol_type="unassigned DNA"
/db_xref="taxon:197"
/notice="11.5 kb PCR product from C. jejuni OH4384 including
LOS biosynthesis locus"

ORIGIN
Query Match          95.6%; Score 837.6; DB 2; Length 11474;
Best Local Similarity 97.3%; Pred. No. 1.3e-220;

```



Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db |||||  
Qy 6048 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6107  
Db |||||

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db |||||

Qy 6108 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167  
Db |||||

Qy 121 CTGCTGTAATGCAAGAGAGTGTCTTACACCCCTAAATTTCTCTTTGGACATACTAC 180  
Db |||||

Qy 6168 CTGCTGTAATGCAAGAGAGTGTCTTACACCCCTAAATTTCTCTTTGGACATACTAC 6227  
Db |||||

Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGTTAAAT 240  
Db |||||

Qy 6228 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGTTAAAT 6287  
Db |||||

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTGTAACCTTTTACGATTTATTTTCT 300  
Db |||||

Qy 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTGTAACCTTTTACGATTTATTTTCT 6347  
Db |||||

Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360  
Db |||||

Qy 6348 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 6407  
Db |||||

Qy 361 TTTTACGAAATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGTTAAAT 420  
Db |||||

Qy 6408 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 6467  
Db |||||

Qy 361 TTTTACGAAATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGTTAAAT 420  
Db |||||

Qy 6408 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 6467  
Db |||||

Qy 421 CTAGTGTATTAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGATTA 6887  
Db |||||

Qy 841 CCTAGTGTATTAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGATTA 876  
Db |||||

Qy 6888 CCTAGTGTATTAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGATTA 6923  
Db |||||

## RESULT 40

BD249790 11474 bp DNA linear PAT 17-JUL-2003  
LOCUS  
DEFINITION Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and ganglioside mimics.  
ACCESSION BD249790  
VERSION BD249790.1 GI:33059560  
KEYWORDS JP 2002535992-A/1.  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
Campylobacteraceae; Campylobacter.

REFERENCE  
AUTHORS 1 (bases 1 to 11474)  
TITLE Gilbert, M. and Wakarchuk, W.W.  
JOURNAL Campylobacter glycosyltransferases for biosynthesis of  
COMMENT gangliosides and ganglioside mimics  
PATENT: JP 2002535992-A 1 29-OCT-2002;  
NATIONAL RESEARCH COUNCIL OF CANADA  
OS Campylobacter jejuni  
PN JP 2002535992-A/1  
PD 29-OCT-2002  
PF 01-FEB-2000 JP 2000597438  
PI 01-FEB-1999 US 60/118213.31-JAN-2000 US 06/495406 PI  
MICHEL GILBERT, WARREN W WAKARCHUK  
PC C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC  
Campylobacter glycosyltransferases for biosynthesis of CC  
gangliosides and  
CC ganglioside mimics  
FH Key Location/Qualifiers  
FT source 1..11474  
FT /organism='Campylobacter jejuni'.  
FEATURES  
source  
1..11474  
/organism='Campylobacter jejuni'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:197'

ORIGIN  
Query Match 95.6%; Score 837.6; DB 2; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.3e-220;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6107  
Qy 61 CTACCAATGATTTTGTGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 6108 CTACCAATGATTTTGTGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167  
Qy 121 CTGCTGTAATGCAAGAGAGTGTCTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
Db 6168 CTGCTGTAATGCAAGAGAGTGTCTTACACCCCTAAATTTCTCTTTTGAACAAATACTAC 6227  
Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGTTCTAAT 240  
Db 6228 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGTTCTAAT 6287  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACACTTTTTCGATTTATTTTCTCT 300  
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACACTTTTTCGATTTATTTTCTCT 6347  
Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
Db 6348 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 6407  
Qy 361 TTTTACGAAATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGAGTAGCC 420  
Db 6408 TTTTACGAAATTTTAAATCCAAATCAAGATATGAGACCGCACTAAATATGTTGAGTAGCC 6467  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACACTAGCCCTCGATTTTAA 540  
Db 6528 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACACTAGCCCTCGATTTTAA 6587  
Qy 541 GATGCTCGCATATATCGGACATAGTAAATAACAGATATTAAGCTTTTGAATTTCTA 600  
Db 6588 GATAAATTCACACTATATCGGACATAGTAAATAACAGATATTAAGCTTTTGAATTTCTA 6647  
Qy 601 GAAAAAATTTTCAAAATTAAGATATTTGCTTATGCTCTAATAGCTTTTGAAGCAATTTT 660  
Db 6648 GAAAAAATTTTCAAAATTAAGATATTTGCTTATGCTCTAATAGCTTTTGAAGCAATTTT 6707  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 6767  
Qy 721 AAAGATATCTATACCTTCTAGTGAGCTTTATCGGAAATTTTTCAGAAAAATTTTAAATTTT 780  
Db 6768 AAAGATATCTATACCTTCTAGTGAGCTTTATCGGAAATTTTTCAGAAAAATTTTAAATTTT 6827  
Qy 781 AAAAAATTAAGATTAAGAAATGTTTATTAAGATTTGATTAAGATTTTAAAGATTTTAAAGATTA 840  
Db 6828 AAAAAATTAAGATTAAGAAATGTTTATTAAGATTTGATTAAGATTTTAAAGATTTTAAAGATTA 6887  
Qy 841 CCTAGTGTATTAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGATTA 876  
Db 6888 CCTAGTGTATTAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGAGATTTTAAAGATTA 6923



```

|||||
6648 GAAAAAACTTACAAAATAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAAAATTTT 6707
Qy
661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 720
Db
6708 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 6767
Qy
721 AAAGATATATCATACCTTCTAGTGAGCGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780
Db
6768 AAAGATATATCATACCTTCTAGTGAGCGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 6827
Qy
781 AAAAAATAAAAAATTAAGAAAAATGTTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Db
6828 AAAAAATAAAAAATTAAGAAAAATTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887
Qy
841 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 876
Db
6888 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 6923

RESULT 41
AR271699
LOCUS AR271699 11474 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6503744.
ACCESSION AR271699
VERSION AR271699.1 GI:29703244
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 11474)
Gilbert, M. and Wakarchuk, W. W.
Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
Patent: US 6503744-A 1 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1. 11474
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.6%; Score 837.6; DB 2; Length 11474;
Best Local Similarity 97.3%; Pred. No. 1.3e-220;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 6048 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 6107
Qy 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120
Db 6108 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 6167
Qy 121 CTGGTAAAAATGCAAGCGAGTGTTTTACACCCCTAATTTCTTCTTTTGAGCAATACACT 180
Db 6168 CTGGTAAAAATGCAAGCGAGTGTTTTACAACTCTATTCTTTTGTGAAACAATACTAC 6227
Qy 181 ACTTTAAACATTTTAAATCCAAAAATCAAGATATAGACCCGAACTTAATTTATGTTCTAAT 240
Db 6228 ACTTTAAACATTTTAAATCCAAAAATCAAGATATAGACCCGAACTTAATTTATGTTCTAAT 6287
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATATTTCCT 300
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATATTTCCT 6347
Qy 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360
Db 6348 GATGCTCATTTGGGATATGATTTTTTTTCAACAACTTAAAGAAATTTTAAATGCTTTATTTAA 6407
Qy 361 TTTCACAAATTTATTTCATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 420
|||||

```

```

Db 6408 TTTCACGAAATTTATTTCATCAATCAAGAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 6467
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 6527
Qy 481 TCTTATGCTTTTGATATACCAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 6528 TCTTATGCTTTTGATATACCAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 6587
Qy 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 6588 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6647
Qy 601 GAAAAAATCAAAAAATTAAGTATATGCTTCTTAAAGTCTTTTACCAATTTT 660
Db 6648 GAAAAAATCAAAAAATTAAGTATATGCTTCTTAAAGTCTTTTACCAATTTT 6707
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 720
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 6767
Qy 721 AAAGATATATCATACCTTCTAGTGAGCGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780
Db 6768 AAAGATATATCATACCTTCTAGTGAGCGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 6827
Qy 781 AAAAAATAAAAAATTAAGAAAAATGTTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 6828 AAAAAATAAAAAATTAAGAAAAATTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 6887
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 876
Db 6888 CCTAGTATATAAGCAATTTTCAAGGAAAAATAA 6923

RESULT 42
AR481781
LOCUS AR481781 11474 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 1 from patent US 6699705.
ACCESSION AR481781
VERSION AR481781.1 GI:47243416
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 11474)
Gilbert, M. and Wakarchuk, W. W.
Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
Patent: US 6699705-A 1 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1. 11474
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.6%; Score 837.6; DB 2; Length 11474;
Best Local Similarity 97.3%; Pred. No. 1.3e-220;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 6048 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 6107
Qy 61 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120
Db 6108 CTACCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 6167
Qy 121 CTGGTAAAAATGCAAGCGAGTGTTTTACACCCCTAATTTCTTCTTTTGAGCAATACACT 180
Db 6168 CTGGTAAAAATGCAAGCGAGTGTTTTACAACTCTATTCTTTTGTGAAACAATACTAC 6227

```

```
Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAATATGTTGTTCTAAT 240
|
|
|
Db 6228 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAATATGTTGTTCTAAT 6287
|
|
|
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATATTTTCCCT 300
|
|
|
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATATTTTCCCT 6347
|
|
|
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTATTTTAA 360
|
|
|
Db 6348 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAAGCTTATTTTAA 6407
|
|
|
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
|
|
|
Db 6408 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 6467
|
|
|
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
|
|
|
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 6527
|
|
|
Qy 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
|
|
|
Db 6528 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6587
|
|
|
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTAGAAATTTCTA 600
|
|
|
Db 6588 GATAAATTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTAGAAATTTCTA 6647
|
|
|
Qy 601 GAAAAAATAAAAATTTAAAGAAATGTTTATACAAAGTTGATAAGAAAGCTTATTAAGATTA 840
|
|
|
Db 6828 AAAAAATAAAAATTTAAAGAAATGTTTATACAAAGTTGATAAGAAAGCTTATTAAGATTA 6887
|
|
|
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876
|
|
|
Db 6888 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 6923
|
|
|
```

```
RESULT 43
AR527380
LOCUS AR527380 11474 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6723545.
ACCESSION AR527380
VERSION AR527380.1 GI:53914401
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
```

```
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 1 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1. .11474
/mol_type="genomic DNA"
ORIGIN
Query Match 95.6%; Score 837.6; DB 2; Length 11474;
```

```
Best Local Similarity 97.3%; Pred. No. 1.3e-220;
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
|
|
|
Db 6048 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGA 6107
|
|
|
Qy 61 CTACCAAAATGATTTTGTATGTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
|
|
|
Db 6108 CTACCAAAATGATTTTGTATGTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 6167
|
|
|
Qy 121 CTTGTAAAAATGCAAGAGGTTTACACCCCTAAATTTCTTTTGTGAGCAATACTAT 180
|
|
|
Db 6168 CTTGTAAAAATGCAAGAGGTTTATCAATCTCTCTTTTGTGAGCAATACTAT 6227
|
|
|
Qy 181 ACTTTAAAAATTAATCCAAATCAAGATATGAGACCGAAGCTTAAATATGTTGTTCTAAT 240
|
|
|
Db 6228 ACTTTAAAAATTAATCCAAATCAAGATATGAGACCGAAGCTTAAATATGTTGTTCTAAT 6287
|
|
|
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300
|
|
|
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 6347
|
|
|
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
|
|
|
Db 6348 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 6407
|
|
|
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
|
|
|
Db 6408 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 6467
|
|
|
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
|
|
|
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 6527
|
|
|
Qy 481 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
|
|
|
Db 6528 TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6587
|
|
|
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTAGAAATTTCTA 600
|
|
|
Db 6588 GATAAATTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTAGAAATTTCTA 6647
|
|
|
Qy 601 GAAAAAATAAAAATTTAAAGAAATGTTTATACAAAGTTGATAAGAAAGCTTATTAAGATTA 660
|
|
|
Db 6648 GAAAAAATAAAAATTTAAAGAAATGTTTATACAAAGTTGATAAGAAAGCTTATTAAGATTA 6707
|
|
|
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAATCAAGGAAAAATAA 876
|
|
|
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAATCAAGGAAAAATAA 6767
|
|
|
Qy 721 AAGATATATCATCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780
|
|
|
Db 6768 AAGATATATCATCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTT 6827
|
|
|
Qy 781 AAAAAATAAAAATTTAAAGAAATGTTTATACAAAGTTGATAAGAAAGCTTATTAAGATTA 840
|
|
|
Db 6828 AAAAAATAAAAATTTAAAGAAATGTTTATACAAAGTTGATAAGAAAGCTTATTAAGATTA 6887
|
|
|
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876
|
|
|
Db 6888 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 6923
|
|
|
```

```
RESULT 44
AR609659
LOCUS AR609659 11474 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 1 from patent US 6825019.
ACCESSION AR609659
VERSION AR609659.1 GI:56664959
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
```

Unclassified.  
1 (bases 1 to 11474)  
Gilbert,M. and Wakarchuk,W.W.  
Polypeptides having .beta.-1,3-galactosyl transferase activity  
Patent: US 6825019-A 1 30-NOV-2004;  
National Research Council of Canada; Ottawa;  
CAX;  
Location/Qualifiers  
1. .11474  
/organism="unknown"  
/mol\_type="genomic DNA"

Query Match 95.6%; Score 837.6; DB 2; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.3e-220;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60  
Db 6048 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATCAAGA 6107

Qy 61 CTACCAAAATGATTGATGATTAGATGAATCAATTTTATTGTAAGATAAATACTAT 120  
Db 6108 CTACCAAAATGATTGATGATTAGATGAATCAATTTTATTGTAAGATAAATACTAT 6167

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTTCTTTGGACATACTAC 180  
Db 6168 CTTGGTAAAAAATGCAAGCAGTGTTTTACAATCCTATTCTTTTGTGAACAATACTAC 6227

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAAATATGTTGTTCTAAT 240  
Db 6228 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAAATATGTTGTTCTAAT 6287

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAGAACTTTTACGATTAATTTTCCT 300  
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAGAACTTTTACGATTAATTTTCCT 6347

Qy 301 GATGCTCAITTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360  
Db 6348 GATGCTCAITTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 6407

Qy 361 TTTTACGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 420  
Db 6408 TTTTACGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 6467

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 6527

Qy 481 TCTTATGCTTTTGATACCAAACTTAAAGAAATTTTAAACAACTTAAAGAAATTTTAAAT 540  
Db 6528 TCTTATGCTTTTGATACCAAACTTAAAGAAATTTTAAACAACTTAAAGAAATTTTAAAT 6587

Qy 541 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAAGATTTCTA 600  
Db 6588 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAAGATTTCTA 6647

Qy 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660  
Db 6648 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 6697

Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATATCATATACAGAAAAATAATCACTAC 720  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATATCATATACAGAAAAATAATCACTAC 6767

Qy 721 AAAGATATCTATACCTCTAGTAGGCTTATCGGAAATTTTCAAAATTTTCAAAATTTTCAAAATTT 780  
Db 6768 AAAGATATCTATACCTCTAGTAGGCTTATCGGAAATTTTCAAAATTTTCAAAATTTTCAAAATTT 6827

Qy 781 AAAAAAAATAAAAAATAAAAGAAATTTTATTTTACAAAGTTGATAAAGATCTATTAAGATTA 840  
Db 6828 AAAAAAAATAAAAAATAAAAGAAATTTTATTTTACAAAGTTGATAAAGATCTATTAAGATTA 6887

Qy 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAAATAA 876  
Db 6888 CCTAGTGATATAAAGCATTATTTCAAAGGAAAAATAA 6923

RESULT 45  
AR689937  
LOCUS AR689937 11474 bp DNA linear PAT 12-SEP-2005  
DEFINITION Sequence 1 from patent US 6905867.  
ACCESSION AR689937  
VERSION AR689937.1 GI:74471945  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 11474)  
AUTHORS Gilbert,M. and Wakarchuk,W.W.  
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity  
JOURNAL Patent: US 6905867-A 1 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
Location/Qualifiers  
1. .11474  
/organism="unknown"  
/mol\_type="genomic DNA"

Query Match 95.6%; Score 837.6; DB 2; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.3e-220;  
Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60  
Db 6048 ATGAAAAAGTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGA 6107

Qy 61 CTACCAAAATGATTGATGATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120  
Db 6108 CTACCAAAATGATTGATGATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 6167

Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180  
Db 6168 CTTGGTAAAAAATGCAAGCAGTATTTTACAATCCTATTCTTTTTCGAACAATACTAC 6227

Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAAATATGTTGTTCTAAT 240  
Db 6228 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAAATATGTTGTTCTAAT 6287

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAGAACTTTTACGATTAATTTTCCT 300  
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAGAACTTTTACGATTAATTTTCCT 6347

Qy 301 GATGCTCAITTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360  
Db 6348 GATGCTCAITTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 6407

Qy 361 TTTTACGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 420  
Db 6408 TTTTACGAAATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 6467

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 6527

Qy 481 TCTTATGCTTTTGATACCAAACTTAAAGAAATTTTAAACAACTTAAAGAAATTTTAAAT 540  
Db 6528 TCTTATGCTTTTGATACCAAACTTAAAGAAATTTTAAACAACTTAAAGAAATTTTAAAT 6587

Qy 541 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAAGATTTCTA 600  
Db 6588 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAAGATTTCTA 6647

Qy 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660  
Db 6648 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 6697

Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATATCATATACAGAAAAATAATCACTAC 720  
Db 6708 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATATCATATACAGAAAAATAATCACTAC 6767

Qy 721 AAAGATATCTATACCTCTAGTAGGCTTATCGGAAATTTTCAAAATTTTCAAAATTTTCAAAATTT 780  
Db 6768 AAAGATATCTATACCTCTAGTAGGCTTATCGGAAATTTTCAAAATTTTCAAAATTTTCAAAATTT 6827

Qy 781 AAAAAAAATAAAAAATAAAAGAAATTTTATTTTACAAAGTTGATAAAGATCTATTAAGATTA 840  
Db 6828 AAAAAAAATAAAAAATAAAAGAAATTTTATTTTACAAAGTTGATAAAGATCTATTAAGATTA 6887

6648 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 6707  
661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAAATCTACACT 720  
6708 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAAATCTACACT 6767  
721 AAAGATATCTACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT 780  
6768 AAAGATATCTACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT 6827  
781 AAAAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
6828 AAAAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 6887  
841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
6888 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 6923  
RESULT 46  
AR691833  
LOCUS AR691833 11474 bp DNA linear PAT 13-SEP-2005  
DEFINITION Sequence 1 from patent US 6911337.  
ACCESSION AR691833  
VERSION AR691833.1 GI:74479544  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 11474)  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Nucleic acids encoding .beta.-1,4-GalNAc transferase  
JOURNAL Patent: US 6911337-A 1 28-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES  
source Location/Qualifiers  
1. .11474  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 95.6%; Score 837.6; DB 2; Length 11474;  
Best Local Similarity 97.3%; Pred. No. 1.3e-220;  
Matches 852; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
QY 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60  
DB 6048 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGA 6107  
QY 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATPACTAT 120  
DB 6108 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATPACTAT 6167  
QY 121 CTTGGTAAAAATCAAGACGAGTGTTTTACACCCCTAATTTCTCTTTGACCAATACTAC 180  
DB 6168 CTTGGTAAAAATCAAGACGAGTGTTTTACAACTCCTATTCTTTTTTTTGAACAATACTAC 6227  
QY 181 ACTTTAAACATTTTAAATCCAAAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240  
DB 6228 ACTTTAAACATTTTAAATCCAAAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 6287  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTATTTTCCCT 300  
DB 6288 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCCT 6347  
QY 301 GATGCTCAATTTGGGATATGATTTTAAAACTTAAAGAAATTTAAATGCTTTATTTAAA 360  
DB 6348 GATGCTCAATTTGGGATATGATTTTAAAACTTAAAGAAATTTTAAATGCTTTATTTAAA 6407  
QY 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
DB 6408 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 6467

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480  
DB 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCCGGAAATTTGATTTTATCAAAATGGGTCA 6527  
QY 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540  
DB 6528 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 6587  
QY 541 GATCCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600  
DB 6588 GATAATTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 6647  
QY 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660  
DB 6648 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 6707  
QY 661 ATAGAACTAGCGCCCAAAATTTAAATTTTCAATTTTATCATACAGAAAAATAAATCTACACT 720  
DB 6708 ATAGAACTAGCGCCCAAAATTTAAATTTTCAATTTTATCATACAGAAAAATAAATCTACACT 6767  
QY 721 AAAGATATCTACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT 780  
DB 6768 AAAGATATCTACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT 6827  
QY 781 AAAAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
DB 6828 AAAAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 6887  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
DB 6888 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 6923  
RESULT 47  
AF130984  
LOCUS AF130984 11474 bp DNA linear BCT 09-FEB-2000  
DEFINITION Campylobacter jejuni lipooligosaccharide biosynthesis locus,  
partial sequence.  
ACCESSION AF130984  
VERSION AF130984.1 GI:6940827  
KEYWORDS  
SOURCE  
ORGANISM Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE 1 (bases 1 to 11474)  
AUTHORS Gilbert, M., Brisson, J.R., Karwaski, M.F., Michniewicz, J.,  
Cunningham, A.M., Wu, Y., Young, N.M. and Wakarchuk, W.W.  
TITLE Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.  
Identification of the glycosyltransferase genes, enzymatic  
synthesis of model compounds, and characterization of nanomole  
amounts by 600-mhz (1)h and (13)c NMR analysis  
J. Biol. Chem. 275 (6), 3896-3906 (2000)  
JOURNAL PUBMED 10660542  
REFERENCE 2 (bases 1 to 11474)  
AUTHORS Gilbert, M., Michniewicz, J. and Wakarchuk, W.W.  
TITLE Direct Submission  
JOURNAL Submitted (24-FEB-1999) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
OR6, Canada  
FEATURES  
source Location/Qualifiers  
1. .11474  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="OH4384"  
/db\_xref="taxon:197"  
<1. .360  
/note="Helicobacter pylori RfaC homolog"  
/codon\_start=1  
/transl\_table=11  
/product="heptosyltransferase I"  
/protein\_id="AAF31765.1"

```
/db_xref="GI:6940828"
/translacion="KEYEPAKEVLNLSGIDETHIELAPKENLELMAFTKMDLIITGN
DSGPTHLAPALNKASITIFGATPSYRNAFQTHINKIDTGCKIQNAKHIDKSDFCITR
1EEDIFKLAUGLNEK"
350. .1237
/notes="Helicobacter pylori Waam homolog"
/codon_start=1
/transl_table=11
/product="lipid A biosynthesis acyltransferase"
/protein_id="AAF31766.1"
/db_xref="GI:6940829"
/translacion="MKNSDRIYLSLYILKFFVTFMPDCILHFLALIVARIAFLHNKK
HRKIINTNLQICFPQTKERDLSIKIYENFPAQFGIDCQONQNTTKEKILNKVNFIN
ENFLIDALAKRPILITTAHVGWNEILSLAYAKGAIISIVGKKLSEVWEILQSOSR
TQPDILIDKGGIRQMLSAKKEALGILTDQCVENESVRLKFNKEVNYQMGASL
IAQSNALLIIPVAYKEGGKFCIEFFKAKDSQASLEELTYQAQSCNEVKKRPMWEY
FFHRRFPASYNIEYKGA"
1234. .2490
/notes="Neisseria meningitidis LgtF homolog"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAF31767.1"
/db_xref="GI:6940830"
/translacion="MNLKQISVIIIVKNABOTLECLNSLKDPEDEILLNNESSDNTL
KIANEKPDPANLIYHNAFIGALKNLALSAYAKNDWILSIDADEVLENECIELKN
LKQEDNIIALSRLNKGWIKACGWPDYVLRIFNKNFTFRNDNLVHESLVLPSNA
KTYLKNGLKHSYQIDSHLIDKQYSSUWAKONTHKSGVLKANLRAFWTFRNYP
LKNGLFYGVKGFTIISVCSALGTFPKYMKLYELQRPKTCALIIITYNOKERLKLVD
SVKNLAPLNEVLIIADGSKEDTARLIEYQKPCPLKHWOEDGFKLSKSRNKTII
KNADSYIIVIDGDMILEKDFIKEHLEFAQKFLQGSRRVLNKKSESEILNKDDYRI
IFNKKDFKSKNSFLAKIFYLSKKR"
2786. .3955
/notes="similar to Streptococcus pneumoniae Cps14j"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAF31768.1"
/db_xref="GI:6940831"
/translacion="MKKIVIPYINVEKYLRECLDSVINQTYNLEILLVNDGSTDE
OLDGNNPYNIKAYKQSOAFNEKDLNFTYPSIDYIIFLDSNWKLNCTIEECVIRM
HSLNIATKEYLTKDKRITLFDKNGGLSSARNIGIEYFSGYKLNKTKQHKENSLEIF
KNVDLWFQDHCITYEONIKNKKHKTMEIFDFKECIIITPEYANRALSVGSRDISPG
WNGMIDFNFLKQLKXFINFINEDIHFGIILPASANKIYIVLSOKLYCLRLRANSIN
HDKKITKANVSEYFKOIYETFGENAKAQNLYKAASRVITLXKLIEFFDKQKONALA
IKETFLPCYAKKALMIKKFKDPLNLEQLVLKIPFIQTQLPYDIWKFQKIKNI"
4025. .5068
/gene="cgtA"
/genes="cgtA"
4025. .5068
/notes="glycosyltransferase"
/experiment="experimental evidence, no additional details
recorded"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAF31769.1"
/db_xref="GI:6940832"
/translacion="MLFQSYFVKIICLFIPIRKIRHKIKKTFLLKNQIDKIDSLPK
KTLVQINKYNNEDILKNAIIGEGHKGYNDEKSDPKSLNPAWAFIRVKNEAITL
KASLESILPAIQGVIGYNDCTGSEBIIILEFKQPSFIPKYPBIEIQONPKSEEN
KLYSYNYVASFIPKDEWLKIDVDHIYDAKLYKSYIPNKNYDVVVSIRVDIHYFN
DNFLCKDNNGTILKEPGDCLINNNYLNKWEVLIDRINNKKATKATQSPSSNIHSE
OLAKYKRIILFHTLNNVHFFLKKHQAQDIKYNNWISIEBFFKPYLQNIHNKIEPSMI
SKETLKKIFLTLN"
complement(5054. .5959)
/gene="cgtB"
/genes="cgtB"
complement(5054. .5959)
/notes="glycosyltransferase"
/experiment="experimental evidence, no additional details
recorded"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAF31770.1"
/db_xref="GI:6940833"
/translacion="MFKIILPYINVEQYIARAIESCINQTFKDIELIIVVDDCGNDN
SINTAKYSKDKRIKIHNKNIIGLLRAREYGVKANSPIYIMFLIDDDYELNACEE
CKILDEQDEVLDVFNAINVESNYISYKDFNSGYSKFEVKKIIAKKNLYWTWNG
KLIRKLYLEAFASLRLKDVKNMAEDVLLYVPMLSQAQKIAYNCLNLYHYVPNNNS
ICNTKNEVLKNNIQELQVLNLRQNYILNKYCSVLVLYIKLYIYIYIKRTKLM
VTLKANILITLKLFLPKYKFLKQC"
6048. .6923
/genes="cst-II"
/genes="cst-II"
/experiment="experimental evidence, no additional details
recorded"
/notes="bifunctional sialyltransferase"
/codon_start=1
/transl_table=11
/product="alpha-2,3/8-sialyltransferase"
/protein_id="AAF31771.1"
/db_xref="GI:6940834"
/translacion="MKKVIITAGNPSLKEIDYSRPNDFVPRCNOQFVFKYVYLGKK
CKAYVNPILFEQYITLKHIIQOEYETELIMCSYNOAHLENENFVKTFDYDFPDA
HLGVDYFKQKDFNAYFKFHEIYFNQRTISGYMCAVAIYALGYKEIYLSGIDFYQNGS
SYAFDTQKLLKLAPNPNDSNVHIGHSKNTDKALFEKTYKIKLYCLCPNSLLA
NFELAPNLNSNFIQEKNTYTKDILIPSEASSEYKFSKNINFKIKENIYYKLKD
LLRUPSDIKHYFKGK"
6924. .7964
/notes="Neisseria meningitidis SiaC homolog"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAF31772.1"
/db_xref="GI:6940835"
/translacion="MKEIKIQNIISSEKAPLVPEIGINHNGSLELAKIMVDAFST
GAKIKHQTHIVEDEMSKAAKVIIPGNAKISIEIMOKALDYKDELALKEYTEKLG
VYLSTPFRAGANLEDMGVSAGFKIGSECNNYPLIKHIAFKPMIVSTGMSIESI
KPTVKILLDNEIPFVLMHTNLVPPHNLVRLNMLELKEBSCMVGLSDYTDNLAC
LGVALGACVLERHFTDSMHRSGPDVCSMDTQALKELIQSEQWAIHRGNESKAA
KQEQYITDFAFASVSVIKDIKKEVLSMDNIVWVRPGLGGISAAEFNLGKALRDI
ENDQLSYEDFA"
8021. .9079
/notes="involved in sialic acid biosynthesis; Neisseria
meningitidis SiaA homolog"
/codon_start=1
/transl_table=11
/product="SiaA-like protein"
/protein_id="AAF31773.1"
/db_xref="GI:6940836"
/translacion="MYRVQNSSEFLYIFATGMHLSKNFGYTVKELYKNGFKNIYEFI
NDYKFTSDKALATTIDGDSRYNELKPDLIIVHGDRIEPLAAALVAGLNNILVAHIE
GGEISGTIDDSLRHAIKLAHLHLVNDFAKRLMQLGDEKSIPIIGSPDLELNDN
KISLNAKYYVDINVENYALLMHPVTEITSIKNQADNLKALIQSNKNYIYIYFNN
DLGELILQSYEELKNNPRFKLPSPLEFYITLLKNADFIIGNSSCILKEALYLKTA
GILVSGRQNGRLGHNTELKVANSDEILKALNTIHKQDLFSAKLEILDSSSKLFFEYL
QSSEFFKLNTOKTVKFDIK"
9076. .9741
/notes="Haemophilus ducreyi NeuA homolog"
/codon_start=1
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AAF31774.1"
/db_xref="GI:6940837"
/translacion="MSLAIIPARGSGKIGKNNVLNNKPIYIYTIKAALNTKISK
VVSSDSDEILNVAKSONVDILRPISLAQDNTTSDKVLHLKALFKVDYDFLOPT
VWSSDDEILNVAKSONVDILRPISLAQDNTTSDKVLHLKALFKVDYDFLOPT
SPLTNTHIDEAFNLYKSNANALISVECDNKLKAFVCEVGLAGI/CNDEYPRMP
RQKLPKTYMSNGAIYILKIKFELNNPSLQSKTKHFLMDESSSLDCLDLKAEQI
WKK"
9729. .10562
/notes="similar to Rhodobacter capsulatus hypothetical
protein encoded by GenBank Accession Number AF010496"
/codon_start=1
```

```

/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAF31775.1"
/db_xref="GI:6940838"
/translation="MEKILKCNKNILMLKQYNIYKYIENPRPRLTKPFIPT
PLENNQLESVAGLIEYCAFKFNILHEMDSFSGFLPHYTKVGRYCSISDGVSM
FNQHPMDRI TASFTYETNHSFINDACQNHKTFPIVNHNPSSSIHTLIQDDVWI
KNDVLLKGIITLGTGCVIGORAVTKDVPYAI VAGI PAKI I KYRFDKTIERLLKIQ
WKYHFDADFYDILNLKINQYLDLLEKIIKKSISYNNPKLYPRDILELKSXKIFNL
F"

```

CDS

```

complement(10554..11366)
/note="similar to Haemophilus influenzae hypothetical
protein HI0868 encoded by GenBank Accession Number U32768"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAF31776.1"

```

Query Match	95.6%;	Score 837.6;	DB 15;	Length 11474;
Best Local Similarity	97.3%;	Pred. No. 1.3e-220;	Mismatches 0;	Gaps 0;
Matches 852;	Conservative 0;	Mismatches 24;	Indels 0;	Gaps 0;
QY	1	ATGAAAAAGATTATTATTCTGGAAATGGACCAAGCTTTAAAGAAATTCATTATTCAGG	60	
DB	6048	ATGAAAAAGATTATTATTCTGGAATGGACCAAGCTTTAAAGAAATTCATTATTCAGA	6107	
QY	61	CTACCAATGATTTGATGTAATAGATGTAATCAATTTATTTTGAAGATAAATACTAT	120	
DB	6108	CTACCAATGATTTGATGTAATAGATGTAATCAATTTATTTTGAAGATAAATACTAT	6167	
QY	121	CTTGCTAAATAATGCAAGCAGTGTTTACACCCCTTAATTTCTCTTGAGCAATACTAC	180	
DB	6168	CTTGCTAAATAATGCAAGCAGTATTTTACAACTCTTCTTTTGTGAACTACTAC	6227	
QY	181	ACTTTAAACATTTAATCAAAATCAAGAATATGACCGAACTAATATGTTCTTAAT	240	
DB	6228	ACTTTAAACATTTAATCAAAATCAAGATATGACCGAACTAATATGTTCTTAAT	6287	
QY	241	TACAAACCAAGCTCACTAGAAAATGAAAATTTGTGAAAACCTTTTACGATTATTTCT	300	
DB	6288	TACAAACCAAGCTCACTAGAAAATGAAAATTTGTGAAAACCTTTTACGATTATTTCT	6347	
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAATGCTTATTTTAA	360	
DB	6348	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAATGCTTATTTTAA	6407	
QY	361	TTTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGC	420	
DB	6408	TTTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGC	6467	
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA	480	
DB	6468	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA	6527	
QY	481	TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT	540	
DB	6528	TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT	6587	
QY	541	GATCGCTCGCACTATATCGGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA	600	
DB	6588	GATAATTCACACTATATCGGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA	6647	
QY	601	GAATAATTCACACTATATCGGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTT	660	
DB	6648	GAATAATTCACACTATATCGGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTT	6707	
QY	661	ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATAACAAGAAAAAATACTACACT	720	
DB	6708	ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATAACAAGAAAAAATACTACACT	6767	
QY	721	AAAGATATACATACCTTCTAGTCAGGCTTATGGAATTTTCAAAAATATTAATTTT	780	
DB	6768	AAAGATATACATACCTTCTAGTCAGGCTTATGGAATTTTCAAAAATATTAATTTT	6827	

QY	781	AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTCATAAAGATCTATTAAAGATTA	840	
DB	6828	AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTCATAAAGATCTATTAAAGATTA	6887	
QY	841	CCTAGTGATATATAAGCAATTTTCAAAGGAAAAATAA	876	
DB	6888	CCTAGTGATATATAAGCAATTTTCAAAGGAAAAATAA	6923	
RESULT 48				
AY297047		12576 bp	DNA	linear
LOCUS				BCT 28-JUL-2003
DEFINITION		Campylobacter jejuni HB93-13 wlaII LOS biosynthesis gene cluster,		
		partial sequence.		
ACCESSION	AY297047			
VERSION	AY297047.1	GI:33151090		
KEYWORDS				
SOURCE				
ORGANISM				
		Campylobacter jejuni		
		Campylobacter jejuni		
		Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;		
		Campylobacteraceae; Campylobacter.		
REFERENCE				
AUTHORS		1 (bases 1 to 12576)		
TITLE		Nawagamuwa,V., Coloe,P.J. and Fry,B.N.		
JOURNAL		The wlaII LOS biosynthesis locus of C. jejuni HB93-13 (HS:19)		
REFERENCE		2 (bases 1 to 12576)		
AUTHORS		Nawagamuwa,V., Coloe,P.J. and Fry,B.N.		
TITLE		Direct Submission		
JOURNAL		Submitted (12-MAY-2003) Biotechnology & Environmental Biology,		
		Royal Melbourne Institute of Technology University, Plenty Road,		
		Bundoora, VIC 3083, Australia		
FEATURES				
		Location/Qualifiers		
		1..12576		
		/organism="Campylobacter jejuni"		
		/mol_type="genomic DNA"		
		/strain="HB93-13"		
		/serotype="HS:19"		
		/db_xref="taxon:197"		
		<1..>12576		
		/note="contains wlaII LOS biosynthesis gene cluster"		
misc_feature				
ORIGIN				
Query Match	95.6%;	Score 837.6;	DB 15;	Length 12576;
Best Local Similarity	97.3%;	Pred. No. 1.2e-220;	Mismatches 0;	Gaps 0;
Matches 852;	Conservative 0;	Mismatches 24;	Indels 0;	Gaps 0;
QY	1	ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG	60	
DB	6063	ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATTCATTATTCAGA	6122	
QY	61	CTACCAATGATTTTGATGTAATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120	
DB	6123	CTACCAATGATTTTGATGTAATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	6182	
QY	121	CTTGCTAAATAATGCAAGCAGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC	180	
DB	6183	CTTGCTAAATAATGCAAGCAGTATTTTACATCTCTTCTTTTGTGAACTACTAC	6242	
QY	181	ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT	240	
DB	6243	ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTTCTTAAT	6302	
QY	241	TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTTATTTCCCT	300	
DB	6303	TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTTATTTCCCT	6362	
QY	301	GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	360	
DB	6363	GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	6422	
QY	361	TTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC	420	





KEYWORDS JP 2002535992-A/2.  
SOURCE Campylobacter jejuni  
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert, M. and Wakarchuk, W. W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and ganglioside mimics  
JOURNAL Patent: JP 2002535992-A 2 29-OCT-2002;  
NATIONAL RESEARCH COUNCIL OF CANADA  
COMMENT OS Campylobacter jejuni  
PN JP 2002535992-A/2  
PD 29-OCT-2002  
PF 01-FEB-2000 JP 2000597438  
PR 01-FEB-1999 US 60/118213, 31-JAN-2000 US 06/495406 PI  
MICHEL GILBERT, WARREN W WAKARCHUK  
PC C12N15/09, C12N1/21, C12N9/88, C12N15/00 CC  
Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and  
CC ganglioside mimics  
FH Key Location/Qualifiers  
FT source 1..876  
/organism='Campylobacter jejuni'.  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:197"

FEATURES  
source  
LOCUS AR271700 876 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 2 from patent US 6503744.  
ACCESSION AR271700  
VERSION AR271700.1 GI:29703245  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert, M. and Wakarchuk, W. W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
JOURNAL Patent: US 6503744-A 2 07-JAN-2003;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source  
Location/Qualifiers  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 95.4%; Score 836; DB 2; Length 876;  
Best Local Similarity 97.1%; Pred. No. 1e-219;  
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAATCGAACATATGAGCCGAATTAATGTTCTTAAT 60  
DB 1 ATGAAAAAGTTATTATTGCTGGAATCGAACATTAATGTTCTTAATGTTCTTAAT 60  
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTTAAAAATGCAAAAGCAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180  
DB 121 CTGTTAAAAATGCAAGGCAGTATTTTACAATCTTCTTTTGAACAACTACTAC 180  
QY 181 ACTTTAAAAATTTAATCCAAATCAAGAATATGAGCCGAATTAATGTTCTTAAT 240  
DB 181 ACTTTAAAAATTTAATCCAAATCAAGAATATGAGCCGAATTAATGTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTCAACAACTTAAAGATTTTAATGCTTTTAAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTCAACAACTTAAAGATTTTAATGCTTTTAAAA 360

Db 541 GATAATTCACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660  
DB 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATACTACTACT 720  
DB 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATACTACTACT 720  
QY 721 AAAGATATATCTATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
DB 721 AAAGATATATCTATACCTTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAAATAAAAAATTAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
DB 781 AAAAAAATAAAAAATTAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATATAAGCATTTATTTCAAAGGAAAAATAA 876  
DB 841 CCTAGTGATATATAAGCATTTATTTCAAAGGAAAAATAA 876



```
Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGACTATATCGGCATAGTAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
Db 541 GATCGCTCGACTATATCGGCATAGTAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
Qy 601 GAAAAAATCTTACAAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTACAAAATTTT 660
Db 601 GAAAAAATCTTACAAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTACAAAATTTT 660
Qy 661 ATAGAACTAGCGCCCAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTACACT 720
Db 661 ATAGAACTAGCGCCCAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTACACT 720
Qy 721 AAAGATATATCATACACTCTTAGTGAGGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 780
Db 721 AAAGATATATCATACACTCTTAGTGAGGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 780
Qy 781 AAAAAAATATAAATTTAAAGAAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTA 840
Db 781 AAAAAAATATAAATTTAAAGAAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876

RESULT 52
AR481782
LOCUS AR481782 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 2 from patent US 6699705.
ACCESSION AR481782
VERSION AR481782.1 GI:47243417
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics
JOURNAL Patent: US 6699705-A 2 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 95.4%; Score 836; DB 2; Length 876;
Best Local Similarity 97.1%; Pred. No. le-219;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTTTCAAGA 60
Qy 61 CTACCAATGATTTTGTGATTTATGATGCTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTGATTTATGATGCTAATCAATTTTATTTTGAAGATAAATACTAT 120
```

```
Qy 121 CTTGGTAAAAATGCAAGCGAGTCTTTTACACCCCTAAATTTCTCTTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCGAGTCTTTTACAAATCTTATCTCTTTTGAACAATACTAC 180
Qy 181 ACTTTAAAAACATTTAAATCCAAAATCAAGATATAGAGCCGAACCTAAATATGTTGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAAATCCAAAATCAAGATATAGAGCCGAACCTAAATATGTTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATTTATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATTTATTTTCT 300
Qy 301 GATCCTCATTTGGGATGATGATTTTAAAAACAACCTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATCCTCATTTGGGATGATGATTTTAAAAACAACCTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 540
Db 481 TCTTATGCTTTTGATACCAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 540
Qy 541 GATCGCTCGACTATATCGGCATAGTAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
Db 541 GATAATTCACACTATATCGGCATAGTAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
Qy 601 GAAAAAATATAAATTTAAAGAAATTTATGCTTATGCTCTAATAGTCTTTTACAAAATTTT 660
Db 601 GAAAAAATATAAATTTAAAGAAATTTATGCTTATGCTCTAATAGTCTTTTACAAAATTTT 660
Qy 661 ATAGAACTAGCGCCCAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTACACT 720
Db 661 ATAGAACTAGCGCCCAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTACACT 720
Qy 721 AAAGATATATCATACACTCTTAGTGAGGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 780
Db 721 AAAGATATATCATACACTCTTAGTGAGGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 780
Qy 781 AAAAAAATATAAATTTAAAGAAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTA 840
Db 781 AAAAAAATATAAATTTAAAGAAATTTTAAATTTTAAATTTTATCATACAGAAATAAATCTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876

RESULT 53
AR527381
LOCUS AR527381 876 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 2 from patent US 6723545.
ACCESSION AR527381
VERSION AR527381.1 GI:53914402
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 2 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
```

RESULT	54
AR609660	
LOCUS	AR609660
DEFINITION	Sequence 2 from patent US 6825019.
ACCESSION	AR609660
VERSION	AR609660.1 GI:56664960
	876 bp DNA linear PAT 15-DEC-2004

```
Db 781 AAAAAAAAAAATAAAGAAAAATATTTATACAGTTGATAAAGATCTATTAAAGTTA 840
Qy 841 CCTAGTGATATAAGCATTTATTCAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTCAGGAAAAATAA 876

RESULT 55
AR689938
LOCUS AR689938 876 bp DNA linear PAT 12-SEP-2005
DEFINITION Sequence 2 from patent US 6905867.
ACCESSION AR689938
VERSION AR689938.1 GI:74471946
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl
transferase activity
JOURNAL Patent: US 6905867-A 2 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876 /organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.4%; Score 836; DB 2; Length 876;
Best Local Similarity 97.1%; Pred. No. 1e-219;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60
Qy 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Qy 121 CTGGTAAAAATGCAAGCGAGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Db 121 CTGGTAAAAATGCAAGCGAGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCC 300
Qy 301 GATGCTCATTTGGGATATGATTTTGTAAAAAATTTTGTAAAAAATTTTACGATTATTTCC 360
Db 301 GATGCTCATTTGGGATATGATTTTGTAAAAAATTTTGTAAAAAATTTTACGATTATTTCC 360
Qy 541 GATGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Db 781 AAAAAAAAAAATAAAGAAAAATATTTATACAGTTGATAAAGATCTATTAAAGTTA 840
Qy 841 CCTAGTGATATAAGCATTTATTCAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTCAGGAAAAATAA 876

RESULT 56
AR691834
LOCUS AR691834 876 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 2 from patent US 6911337.
ACCESSION AR691834
VERSION AR691834.1 GI:74479545
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAC transferase
JOURNAL Patent: US 6911337-A 2 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876 /organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.4%; Score 836; DB 2; Length 876;
Best Local Similarity 97.1%; Pred. No. 1e-219;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60
Qy 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTGGTAAAAATGCAAGCGAGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Db 121 CTGGTAAAAATGCAAGCGAGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCC 300
Qy 301 GATGCTCATTTGGGATATGATTTTGTAAAAAATTTTGTAAAAAATTTTACGATTATTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTGTAAAAAATTTTGTAAAAAATTTTACGATTATTTAA 360
```

```
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600
Qy 601 GAAAAAAGCTTACAAAATAAACTATATTCCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAAGCTTACAAAATAAACTATATTCCTTATGCTCTTAACAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATAACAAGAAAAAATACTACACT 720
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATAACAAGAAAAAATACTACACT 720
Qy 721 AAAGATATATCATACCTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATATCATACCTCTAGTGGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
Qy 781 AAAAAAATAAAAAATTAAGAAAAATCTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAAAAATTAAGAAAAATCTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTATTCAGGAAAAATAA 876
Db 841 CCTAGTGATATAAGCATTTATTCAGGAAAAATAA 876

RESULT 56
AR691834
LOCUS AR691834 876 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 2 from patent US 6911337.
ACCESSION AR691834
VERSION AR691834.1 GI:74479545
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAC transferase
JOURNAL Patent: US 6911337-A 2 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876 /organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 95.4%; Score 836; DB 2; Length 876;
Best Local Similarity 97.1%; Pred. No. 1e-219;
Matches 851; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60
Qy 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTGGTAAAAATGCAAGCGAGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Db 121 CTGGTAAAAATGCAAGCGAGTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCC 300
Qy 301 GATGCTCATTTGGGATATGATTTTGTAAAAAATTTTGTAAAAAATTTTACGATTATTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTGTAAAAAATTTTGTAAAAAATTTTACGATTATTTAA 360
```

Qy 361 TTTTCAGGAAATTTATTTCAATCAAGAATACCTCAGGGGTCTATATGTCGACGTAGCC 420  
Db 361 TTTTCAGGAAATTTATTTCAATCAAGAATACCTCAGGGGTCTATATGTCGACGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600  
Db 541 GATTAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATCTTACAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTATAGCAAAATTT 660  
Db 601 GAAAAAATCTTACAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTATAGCAAAATTT 660  
Qy 661 ATAGAACTAGCCGCAAAATTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCCGCAAAATTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Qy 721 AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAATAATTAATTTT 780  
Db 721 AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAATAATTAATTTT 780  
Qy 781 AAAAAATAAAATTAAGAAATATTTTATACAAAGTTGATAAAAGATCTATTAAGATTA 840  
Db 781 AAAAAATAAAATTAAGAAATATTTTATACAAAGTTGATAAAAGATCTATTAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876

RESULT 57  
AX934429 AX934429 876 bp DNA linear PAT 05-JAN-2004  
LOCUS Sequence 6 from Patent WO02074942.  
DEFINITION AX934429  
ACCESSION AX934429  
VERSION AX934429.1 GI:40641676  
KEYWORDS  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE  
1 Gilbert, M. and Wakarchuk, W. W.  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
Patent: WO 02074942-A 6 26-SEP-2002;  
National Research Council of Canada (CA)  
Location/Qualifiers  
1. 876  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1. 876  
/note="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (catII) from C.  
jejuni serotype O:41"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04168.1"  
/db\_xref="GI:40641677"  
CDS  
alpha-2,3/alpha 2,8-sialyltransferase II (catII) from C.  
jejuni serotype O:41"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04168.1"  
/db\_xref="GI:40641677"  
CKAVYPSLFEQVYTLKHLKILQOEYETELIMCSNFQAHLENFQVFKYDFYPPDA  
HLGVDFFKOLKEFNAPKFKHEIYFNORITSGVMCTVAIALGYKEIYLSGIDFYONGS  
SYAEDTKQNLKLLAPNPNKNDNSHYIGHSKNTDKALEFLEKTEIKLYCLCPNSLLA  
NFIELAPNLSNFNIQEKNNYTKOILLIPSSSEAYGKFTKNINFKIKIKENIYYKLKD  
LLRLPSDIKHVFKG"

ORIGIN  
Query Match 93.6%; Score 820; DB 2; Length 876;  
Best Local Similarity 96.0%; Pred. No. 2.7e-215;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTTATTTATGCTGGAATGACCAAGTTTAAAGAAATTTGATTTATCAAGG 60  
Db 1 ATGAAAAAGTTTATTTATGCTGGAATGACCAAGTTTAAAGAAATTTGATTTATCAAGA 60  
Qy 61 CTACCAATGATTTTATGATGATTAATCAATCAATTTTATTTGAGAGATAATACTAT 120  
Db 61 CTACCAATGATTTTATGATGATTAATCAATCAATTTTATTTGAGAGATAATACTAT 120  
Qy 121 CTTGTGAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGTGAAAAATGCAAGCAGTGTTTTACCAATCTCTAGTCTTTTGTGACCAATACTAC 180  
Qy 181 ACTTTAAAAACATTTTAAATCCAAATCAAGAATATGAGACCGAATTAATTTGTGTTCTAAT 240  
Db 181 ACTTTAAAAACATTTTAAATCCAAATCAAGAATATGAGACCGAATTAATTTGTGTTCTAAT 240  
Qy 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300  
Db 241 TTTAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Qy 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGGCC 420  
Db 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCGAGTAGGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600  
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATCTTACAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTATAGCAAAATTT 660  
Db 601 GAAAAAATCTTACAAAATAAAACTATATGCTTATGCTCTAATAGTCTTTTATAGCAAAATTT 660  
Qy 661 ATAGAACTAGCCGCAAAATTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCCGCAAAATTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Qy 721 AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAATAATTAATTTT 780  
Db 721 AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAATAATTAATTTT 780  
Qy 781 AAAAAATAAAATTAAGAAATATTTTATACAAAGTTGATAAAAGATCTATTAAGATTA 840  
Db 781 AAAAAATAAAATTAAGAAATATTTTATACAAAGTTGATAAAAGATCTATTAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876  
RESULT 58  
BD249793  
LOCUS  
DEFINITION  
Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and ganglioside mimics.  
ACCESSION BD249793  
BD249793 876 bp DNA linear PAT 17-JUL-2003  
Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and ganglioside mimics.  
BD249793

[illegible]

```
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACACTTAAAGAAATCAATGCTTATTTTAAA 360
Qy 361 TTTCAAGAAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTCAAGAAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAGAAATACAGATATAAAGAGCTTTAGAAATTTCTA 600
Db 541 GATCAATTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAGAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAACTTACAAATAAAGCTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAACTTACAAATAAAGCTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAAATAACTTACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAAATAACTTACT 720
Qy 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAAAAATTTTACAAAAATATTAATTTT 780
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAAAAATTTTACAAAAATATTAATTTT 780
Qy 781 AAAAAAATAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTA 840
Db 781 AAAAAAATAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876

RESULT 60
AR481784 LOCUS AR481784 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 6 from patent US 6699705.
ACCESSION AR481784
VERSION AR481784.1 GI:47243419
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics
JOURNAL Patent: US 6699705-A 6 02-MAR-2004; National Research Council of Canada; Ottawa; CAX;
FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"

ORIGIN
Query Match 93.6%; Score 820; DB 2; Length 876;
Best Local Similarity 96.0%; Pred. No. 2.7e-215;
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTATATTTCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTATATTTCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGA 60
Qy 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
```

```
Qy 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTATTTTACAACTCCTAGTCTTTTTTTGGCAATACTAC 180
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATGTGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATGTGTTCTAAT 240
Qy 241 TACACCCAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300
Db 241 TTTAACCAAGCTCATCTAGAAATCAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGAAATTTCAATGCTTATTTAAA 360
Qy 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAGAAATACAGATATAAAGAGCTTTTAGAAATTTCTA 600
Db 541 GATCAATTCACACTATATCGGACATAGTAAAGAAATACAGATATAAAGAGCTTTTAGAAATTTCTA 600
Qy 601 GAAAAAACTTACAAATAAAGCTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAACTTACAAATAAAGCTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAAATAACTTACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAAATAACTTACT 720
Qy 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAAAAATTTTACAAAAATATTAATTTT 780
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAAAAATTTTACAAAAATATTAATTTT 780
Qy 781 AAAAAAATAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTA 840
Db 781 AAAAAAATAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876

RESULT 61
AR527383 LOCUS AR527383 876 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 6 from patent US 6723545.
ACCESSION AR527383
VERSION AR527383.1 GI:53914404
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Polypeptides having .beta.-1,4-GalNAC transferase activity
JOURNAL Patent: US 6723545-A 6 20-APR-2004; National Research Council of Canada; Ottawa; CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
```

/mol\_type="genomic DNA"

ORIGIN

Query Match 93.6%; Score 820; DB 2; Length 876;  
Best Local Similarity 96.0%; Pred. No. 2.7e-215;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGA 60  
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACAACTCTAGTCTTTTTTTTGAACAATACTAC 180  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300  
Db 241 TTAACCAAGCTCATCTAGAAAATCAAAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTCAAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Qy 361 TTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCGAGTAGCC 420  
Db 361 TTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGATCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCATATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Db 541 GATAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600  
Qy 601 GAAAAACTTACAAAATAAAGAAATTTGCTTATGCTCTTAATAGTCTTTTACGCAATTTT 660  
Db 601 GAAAAACTTACGAAAATAAAGCTATTTGTTTATGTCTTAACAGTCTTTTACGCAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTATACT 720  
Qy 721 AAAGATATCTCATACCTTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATCTCATACCTTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780  
Qy 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAGATTA 840  
Db 781 AAAAAATAAAAATTAAGAAATATTTTATACAGTTGATAAAGATCTATTAGATTA 840  
Qy 841 CCTAGTGATATAAGCATTTTTCAAAGGAAAAATAA 876  
Db 841 CCTAGTGATATAAGCATTTTTCAAAGGAAAAATAA 876

RESULT 62  
AR609662  
LOCUS AR609662 876 bp DNA linear PAT 15-DEC-2004  
DEFINITION Sequence 6 from patent US 6825019.  
ACCESSION AR609662

AR609662.1 GI:56664962

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE  
1 (bases 1 to 876)

AUTHORS  
Gilbert,M. and Wakarchuk,W.W.

TITLE  
Polypeptides having .beta.-1,3-galactosyl transferase activity

JOURNAL  
Patent: US 6825019-A 6 30-NOV-2004;

National Research Council of Canada; Ottawa;

CAX;

FEATURES  
Location/Qualifiers

source  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 93.6%; Score 820; DB 2; Length 876;  
Best Local Similarity 96.0%; Pred. No. 2.7e-215;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGA 60  
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120  
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACAACTCTAGTCTTTTTTTTGAACAATACTAC 180  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300  
Db 241 TTTAACCAAGCTCATCTAGAAAATCAAAAATTTTGTAAAACTTTTACGATTTATTTTCCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTCAAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Qy 361 TTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCGAGTAGCC 420  
Db 361 TTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGATCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCATATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Db 541 GATAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600  
Qy 601 GAAAAACTTACAAAATAAAGAAATTTGCTTATGCTCTTAATAGTCTTTTACGCAATTTT 660  
Db 601 GAAAAACTTACGAAAATAAAGCTATTTGTTTATGTCTTAACAGTCTTTTACGCAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTATACT 720  
Qy 721 AAAGATATCTCATACCTTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATCTCATACCTTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATTTAATTTT 780



QY 781 AAAAAAAAAAATAAGAAATGTTTATACAAAGTTGATATAAAGATCTATTAAAGTTA 840  
Db 781 AAAAAAAAAAATAAGAAATGTTTATACAAAGTTGATATAAAGATCTATTAAAGTTA 840  
QY 841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876  
Db 841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876  
RESULT 63  
AR689940  
LOCUS AR689940 876 bp DNA linear PAT 12-SEP-2005  
DEFINITION Sequence 6 from patent US 6905867.  
ACCESSION AR689940  
VERSION AR689940.1 GI:74471948  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity  
JOURNAL Patent: US 6905867-A 6 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES  
source Location/Qualifiers  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 93.6%; Score 820; DB 2; Length 876;  
Best Local Similarity 96.0%; Pred. No. 2.7e-215;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAAATGCACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGCACCAAGTTTAAAGAAATTTGATTATTCAAGA 60  
QY 61 CTACCAAAATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAAAGCAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAAAGCAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATCATGTGTTCTAAT 240  
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATCATGTGTTCTAAT 240  
QY 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCCCT 300  
Db 241 TTTAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
QY 541 GATGCTCGCACTATATCGGCATAGTAAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600

Db 541 GATAATTTCACCTATATCGGCACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
QY 601 GAAAAAACTTTACAAAATAAACTATATTGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAACTTTACAAAATAAAAGCTATATTGTTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAAATAACTATACT 720  
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATAATTTT 780  
QY 781 AAAAAAATAAAATTAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876  
Db 841 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 876  
RESULT 64  
AR691836  
LOCUS AR691836 876 bp DNA linear PAT 13-SEP-2005  
DEFINITION Sequence 6 from patent US 6911337.  
ACCESSION AR691836  
VERSION AR691836.1 GI:74479547  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Nucleic acids encoding .beta.-1,4-GalNAC transferase  
JOURNAL Patent: US 6911337-A 6 28-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES  
source Location/Qualifiers  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 93.6%; Score 820; DB 2; Length 876;  
Best Local Similarity 96.0%; Pred. No. 2.7e-215;  
Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60  
QY 61 CTACCAAAATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAAAGCAGTGTGTTTACACCCCTAAATTTCTTTTGTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAAAGCAGTGTGTTTACACCTCTAGTCTTTTGTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATCATGTGTTCTAAT 240  
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATCATGTGTTCTAAT 240  
QY 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCCCT 300  
Db 241 TTTAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCCCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360





/product="putative  
 beta-1,4-N-acetylgalactosaminyltransferase"  
 /protein\_id="AAK95999.1"  
 /db\_xref="GI:15430481"  
 /translation="MLFQSYFVKIICLPIPRKHKKTKTLLKNQIDKIDSLPK  
 KTLAQINNEDELKANKAIGIEHKGFFNFYDEKSPKSLNPAFIRVKNKSEEN  
 KASLESILPQGVINDCTDGEIEILFCKQPSFPIKYPYIEIQINPKSEEN  
 KLYSVNVASPIKDEWLIKIDVDHIIDAKLKYKSFYIPKYNQDVSVYDIHYFN  
 DNFTLCKDNNGLKPGDCLLINNNLKWKEVLIDRINNWKATKOSFSSNIHSL  
 QKYYKHLRFLPTELNNYHPPFLKQRAQDIYKNWISIEEFKKFYQLQINHKIEFSMI  
 SKETUKKIFLTLF"  
 complement (5044. .5949)  
 /gene="cgtb"  
 complement (5044. .5949)  
 /gene="cgtb"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative beta-1,3-galactosyltransferase"  
 /protein\_id="AAK96000.1"  
 /db\_xref="GI:15430482"  
 /translation="MFKSILPTYNVEQYIARAEISCINQTFKDIILVVDGNDN  
 SINAKYKSKDKRIKIHNENKUGLLRARVEGVKANSFYIMFLDPDDYDELNACEE  
 CIKILDEQVDLVFNFAVESNVIYKFFNSGFYSKFEFKKIIANKNLWTMWG  
 KLIRKLYLEAFASLRLEKDKVKNMAEDVLLYPMLSQAQKIAYNNCLNLYHVPNNNS  
 ICNTKNEVLKNNIQELOLVNLQRQVILNKCYSVLVLIKLYLIQIYKIKRTKLM  
 VTLAKINILILKILFKYKFKLQK"  
 6038. .6913  
 /gene="cstII"  
 6038. .6913  
 /gene="cstII"  
 /codon\_start=1  
 /transl\_table=11  
 /product="alpha-2,3-/alpha-2,8-sialyltransferase"  
 /protein\_id="AAK96001.1"  
 /db\_xref="GI:15430483"  
 /translation="MKWVLIISGNPSLKEIDYSLPNDPVDPRCNOQFEDRYLGGK  
 CKAPYNSPLPEQVYTLKHLIQOEYETELIMCSNFQAHLENNFVKTFVDYPPDA  
 HLGIDFKOLKEFNFAFHEIYFNQRTISGYMCTVAIALGKIYLSGIDFYONGS  
 SYAFDTKOKNLLKAPNFKNDNSYIHSKNTDIKALEFLEKTYEIKLYCLPCNSLLA  
 NYIELAPNLNFIIOEKNNYTKOILIPSEAYGKFTKNIINFKIKIENIYKLIKID  
 LLRLPSDIKYHFKGR"  
 6914. .7954  
 /gene="neuB1"  
 6914. .7954  
 /gene="neuB1"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative sialic acid synthase"  
 /protein\_id="AAK96002.1"  
 /db\_xref="GI:15430484"  
 /translation="MKEIKIQNIISBEKAPLVPEIGINHNHNSLELAKIMVDAFSA  
 GAKIKHOTHIVEDMSKAARKVIPGNAKISYIEIMOKCALNYKDELALKEYTEKILG  
 VYLSPPFRAGANRLIEDMGVAFKIGSGECNNYPLIKHIAFPKPMIVSTGMNSTESI  
 KPTVKILDNIEIPVLMHTNLYPTPHNLVRLNMLELKKSFCWVGJSDHTDNLAC  
 LGVALGACVLERHFTDSMRSGDPDIVCSMDTQALKELIIQSEQWAINRGNNESKAA  
 KQOQVTDFAFASVVSIIKDIRKGEVLSMDNIWVRPGLIGGISASEFENILGKKAURDI  
 ENDQLSYEDFT"  
 7951. .9069  
 /gene="neuC1"  
 7951. .9069  
 /gene="neuC1"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative N-acetylglucosamine-6-phosphate"  
 /protein\_id="AAK96003.1"  
 /db\_xref="GI:15430485"  
 /translation="MKKILIFITGRADYSKIKSLMYRVQNSSEFELYIPATGMHLKSN  
 FGYTVKELYKFNKNIYFINDKYQTDKALATIDGFSRYVNLKPDILVVDHRI  
 EPLAAIIVGALNNILVAHIEGIEISGTIDDSLRHAIISKLAIHILVNDSEFAKRLMQLG  
 EDEKSIPIFGISPDNIEILKISLSEAKKYVDINVKYVALLMFHPEVTFEITSIKNOAD  
 NLVKALIOSNNYIVIPNNDLGPFLILQSVLEELKNRPFKLFPSLRFEYFTLLKNA  
 DFIIGNSSCILKEALYLKTAGILVGSRONRIGNENTLKVNNNSDEILKAINTHKKQ  
 DLFSAKVILNSSLFFBYLLSYGSEFFKLNTQKVPKDIK"

Gene 9066. .9731  
 /gene="neuA1"  
 CDS 9066. .9731  
 /gene="neuA1"  
 /codon\_start=1  
 /transl\_table=11  
 /product="CMP-Neu5Ac synthetase"  
 /protein\_id="AAK96004.1"  
 /db\_xref="GI:15430486"  
 /translation="MSLAIIIPARGSGKIKNKLVLNNKPLIYTTIKAAINTKSISK  
 VVSSDSILNAYKQNDVILKRPISLAQDDTTSKVLLHALKFKYDKYKDVFLQPT  
 SEPLRTNIDEAFNLYKSNANALISVSECNDKILKAFVNCYDGLAGINDYFPFIP  
 RQKLPTKTYMSNGAIYILKIKEFLNPSFLQNTKHFILMDESSSIDICLEDLKKVEQI  
 WKK"  
 9719. .9910  
 /note="orf11; premature translation stop due to deletion  
 in homopolymeric G-tract"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative acetyltransferase"  
 /protein\_id="AAK96005.1"  
 Query Match 93.4%; Score 818.4; DB 15; Length 11455;  
 Best Local Similarity 95.9%; Pred. No. 2.5e-215;  
 Matches 840; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 DB 6038 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6097  
 QY 61 CTACCAATGATTTTCATGATGATTAATCAATTTTATTATTGAGATAAATACTAT 120  
 DB 6098 CTACCAATGATTTTCATGATGATTAATCAATTTTATTATTGAGATAAATACTAT 6157  
 QY 121 CTTGTAAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
 DB 6158 CTTGTAAAAAATGCAAGCAGTGTTTACACCTCTAGTCTTTTGTGAACAATACTAC 6217  
 QY 181 ACTTTAAAAACATTTAATCAAAATCAAGAATATGAGCCGAACTAATATTGTGTCTAAT 240  
 DB 6218 ACTTTAAAAACATTTAATCAAAATCAAGAATATGAGCCGAACTAATATTGTGTCTAAT 6277  
 QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCCT 300  
 DB 6278 TTTAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCCT 6337  
 QY 301 GATGCTCATTGGGATATGATTTTAAAAACAATTTTAAAGAAATTTAATGCTTATTTAAA 360  
 DB 6338 GATGCTCATTGGGATATGATTTTAAAAACAATTTTAAAGAAATTTAATGCTTATTTAAA 6397  
 QY 361 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGCC 420  
 DB 6398 TTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGCC 6457  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
 DB 6458 ATAGCCCTAGGATACAAAGAAATTTCTTCGGGAATTCATTTTATCAAAATGGGTCA 6517  
 QY 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
 DB 6518 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 6577  
 QY 541 GATCGCTCCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
 DB 6578 GATAATTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6637  
 QY 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTTATGCTTCTTCTTAGCAAAATTTT 660  
 DB 6638 GAAAAAACTTACAAAATAAAGCTTATATGCTTATGCTTCTTCTTAGCAAAATTTT 6697  
 QY 661 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAAATAAATACTACT 720  
 DB 6698 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAAATAAATACTACT 6757

QY 721 AAGATATACATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATTTAAATTTT 780  
 Db 6758 AAAGATATATCATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATTTAAATTTT 6817  
 QY 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTGTATAAAGATCTATTAAAGATTA 840  
 Db 6818 AAAAAATAAAAAATTAAGAAAAATTTATTACAGTTGTATAAAGATCTATTAAAGATTA 6877  
 QY 841 CCTAGTGATATAAGCATTTATTTCAAGAGAAATPAA 876  
 Db 6878 CCTAGTGATATAAGCATTTATTTCAAGAGAAATPAA 6913

## RESULT 66

AF401529

## LOCUS

## DEFINITION

AF401529 12388 bp DNA linear BCT 03-JAN-2002  
 Campylobacter jejuni O:23 heptosyltransferase I (waaC) gene,  
 partial cds; lipid A biosynthesis acyltransferase (htrB), putative  
 two-domain glycosyltransferase, putative glycosyltransferase,  
 truncated beta-1,4-N-acetylglactosaminyltransferase (cgta-I),  
 truncated beta-1,3-galactosyltransferase (cgta-II), bifunctional  
 alpha-2,3/2,8-sialyltransferase (cst-II), sialic acid synthase  
 (neuB), putative N-acetylglucosamine-6-phosphate 2-epimerase  
 (neuC), beta-1,4-N-acetylglactosaminyltransferase (cgta-II),  
 CMP-Neu5Ac synthetase (neuA), putative acetyltransferase, and  
 putative glycosyltransferase (waaV) genes, complete cds; and  
 heptosyltransferase II (waaF) gene, partial cds.

## ACCESSION

AF401529

## VERSION

AF401529.1 GI:15718497

## KEYWORDS

## SOURCE

## ORGANISM

Campylobacter jejuni

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

Campylobacteraceae; Campylobacter.

1 (bases 1 to 12388)

Gilbert, M., Karwaski, M.F., Bernatchez, S., Young, N.M., Taboada, E.,

Michniewicz, J., Cunningham, A.M., and Wakarchuk, W.W.

The Genetic Bases for the Variation in the Lipo-oligosaccharide of

the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF

STALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE

J. Biol. Chem. 277 (1), 327-337 (2002)

11689567

2 (bases 1 to 12388)

Gilbert, M., Michniewicz, J., Karwaski, M.-F., Cunningham, A. and

Wakarchuk, W.W.

Direct Submission

Submitted (22-JUL-2001) Institute for Biological Sciences, National

Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A

0R6, Canada

## FEATURES

Location/Qualifiers

1..12388

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strain="ATCC 43449"

/serotype="O:23"

/db\_xref="ATCC:43449"

/db\_xref="taxon:197"

&lt;1..348

/gene="waaC"

&lt;1..348

/gene="waaC"

/codon\_start=1

/transl\_table=11

/product="heptosyltransferase I"

/protein\_id="AAL05998.1"

/db\_xref="GI:15718498"

/translation="FAKEVNLGSDIETHIELAPKFSLEELMAFTKMVDLIIGNDSGP

THLAPALNKAISITIFGTPSYRNAFQTHINKIIDAGKKIQNAKHIDKSDFCITRIEKG

DIFKLAKGLNEK"

338..1225

/gene="htrB"

338..1225

/note="phase-variable gene due to homopolymeric G-tract"

/codon\_start=1

/gene="htrB"  
 /codon\_start=1  
 /transl\_table=11  
 /product="lipid A biosynthesis acyltransferase"  
 /protein\_id="AAL05998.1"  
 /db\_xref="GI:15718499"  
 /translation="WKNSDRILYLSYILKFFVTFMPECILHLFALIVARITFYLNKK  
 HRKIIDNLOICFPQYTKERDKLSLKIYENFAQFGIDCLQONTTKKILNKVNPIN  
 ENFLDALAKRPIIFTAHYGNWEILSLVAAYKAISIVGKIKLSVMVYILQSQR  
 TQDIELDKGGIROMLSALKKERALGILTDQCVENESVRLKFNKEVNYQMGSAL  
 IAQRNALIIPYAYKEDGKFCIEFFKAKOSNASLEELTLTQAOSCEMIKCRPWEY  
 FFHRRFASYNEEIYKDAK"  
 1222..2769  
 /note="orf3"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative two-domain glycosyltransferase"  
 /protein\_id="AAL06000.1"  
 /db\_xref="GI:15718500"  
 /translation="WNLKQISVIIIVKNAEOTLLECLNSLKDPEIILINNESSDNTL  
 NIANEPKDFANLYIYHSAFIFGKALKNLALSAYKNDWILSIDADEVLENECIELKN  
 LELEDNIIALSRNLYKGEWIKACGSPDYVLRIPKNTFRPNLNLVHESLIVPSNV  
 KKIHLKNGLRHAFRDISHLINKQYSSSLWAKQNHKKSGVGLNLRFAWTFRRNYF  
 LKNGFLYQOFTISVCNAGTFFKWKLYELOKQPKTCALIIITTYQKERLKLVD  
 SVKNTLPLNEVLADDCGKEDTARLIEEYQKDFPCPSLKHIMQEDGFAAKRNNAI  
 RVKSEYIILIDGMILEKDFIKNHFRFAKVILOGSRVILNKNESEILKNDVSL  
 AFKSKPKSYKNSFLAKIYIKSIDQKFEFTKEIIGSKTCNMSFFKADFDLQGFN  
 ENFIGWGEDSEFVARFLFNKGIFFRLAKFAIYAHYIHEENSKMLESNHQIYLDLIK  
 NKKISWR"  
 2773..3846  
 /note="orf4"  
 /codon\_start=1  
 /transl\_table=11  
 /product="putative glycosyltransferase"  
 /protein\_id="AAL06001.1"  
 /db\_xref="GI:15718501"  
 /translation="MKKIGVPIYVNBVEYLRECLDSVINQTYTNLEIILVNDGSTDE  
 NSLNAKEYTLKDERFILLFKNGGLSSARNVIEYFSGEYKLNKTOHIEKNSLIEF  
 QLDGNPNYIKAYKSSQAPNNEKDTNFTPNIDYIIFLSDDYWEALNSIECVIRM  
 KNDVLPWHDCTYEDNINKHKKTRMEIFDKECIITPPEYANRALSVGRDISFG  
 WNGMIDFNFLKIKLFIINFINEDIHFGIILFASANKIYVLSQKLYLCRLRANSISN  
 HDKKITKANVSEYFKDIYETFGENAKENYKAAASRVTTALKLIEFFKDKQSENALA  
 IKETFLPCYAKKALMIKNLKLIL"  
 3988..4077  
 /gene="cgta-I"  
 3988..4077  
 /gene="cgta-I"  
 /note="orf5-I; premature translation stop due to a one bp  
 deletion"  
 /codon\_start=1  
 /transl\_table=11  
 /product="truncated  
 beta-1,4-N-acetylglactosaminyltransferase"  
 /protein\_id="AAL06002.1"  
 /db\_xref="GI:15718502"  
 /translation="MLFQSYFKIICLPIPRKIRHKIKKHFY"  
 complement(5637..5912)  
 /gene="cgta"  
 complement(5637..5912)  
 /gene="cgta"  
 /codon\_start=1  
 /transl\_table=11  
 /product="truncated beta-1,3-galactosyltransferase"  
 /protein\_id="AAL06003.1"  
 /db\_xref="GI:15718503"  
 /translation="MFKISIIILTYNVQYIARAJESCINQTFKNIEIIVDDCGSDE  
 SIDTAKEYAKDERIKIHNENLGLLARVEGVKAAGGIYVFRP"  
 6001..6876  
 /gene="cst-II"  
 6001..6876  
 /gene="cst-II"  
 /note="phase-variable gene due to homopolymeric G-tract"  
 /codon\_start=1

CDS

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

```
/transl_table=11
/product="bifunctional alpha-2,3/-2,8-sialyltransferase"
/protein_id="AAL06004.1"
/db_xref="GI:15718504"
/translation="MKKKVILSGNPSLKEIDYSLPNDPFCNQFPEBKYYLGGK
CKAVYNSPFEQYITLKHILIOQOEIEILIMCSNYNQAHLENFVKTFYDPPDA
HLGYDFKQLKFNAPYKFEIYFNQRTSGVYMTVAIALGYKEIYLSGLDIFYDNGG
GYADTKQKLLKAPNFKNDNSHYIGHSKNTDILKALEFLEKTYEIKLYCLCPNSLLA
NFIELAPNLSNFIIQEKNNYTKDILIPSSSEAYGKFTKNINFKKIKIKENIYVYKLD
LLKLPDIKHYPKSK"
6877..7917
/gene="neub1"
CDS
6877..7917
/gene="neub1"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAL06005.1"
/db_xref="GI:15718505"
/translation="MKETKIONIIISBEKAPLVPEIGINHNGSLELAKIMVDAFSA
GAKLIKQTHIVEDEMSKAAKVIIPGNAKISILYELMOKCALDYKDELALKEVTEKLG
VILSPFSRAGANRLEDMGVSAFKIGSGECNNYPLIKHIAAFKPMIVSDTGMNSESIE
KPTVKILLDNIPRLVHTTNLYPTPHNLVRLNAMLEUKBFSCMVGLSDHTDNLAC
LGAVALACVLERHFTDSMRHSGPDIVCSMDTOALKELIIQSEQWAIMRGNESKAA
KQEOVTTIDFASVVSVKDIKKGEVLSMDNIWVRPGLGGISAAEFENILGKKALRDI
ENDTQLSVDEFA"
7914..9038
/gene="neuc1"
CDS
7914..9038
/gene="neuc1"
/codon_start=1
/transl_table=11
/product="putative N-acetylglucosamine-6-phosphate
2-epimerase"
/protein_id="AAL06006.1"
/db_xref="GI:15718506"
/translation="MKKILFITGTADYSKISLMYRVONSSEPELYIPATGMHLSKN
EGYTVKELYKNGPNIEFYNYDKYQTDKALATTIDGFSRYVNLKEDPLIVVHGDR
EPLAAVIGALNLIIVAHIGEIGESTGIDDSRAHSKLAHLHVNDEFKRLMQLG
EDGSIPIFGSDPLELNNKISLNEAKYDIDYNYVALLMFHPVTEITSIKNOAD
NLVALLQSNKNYIVIPYNDLGFELILQSYEELKNRPFKLPSLRPEFYITLLKNA
DFIIGNSCSILKEALYLTAGLILVGSRONRGLNENTLKVANSDEILKALINTIHKQ
DLFSTKLEILDSKLFPEYLLSGBEFFKLSTQVKFDIKHA"
9031..9978
/gene="cgta-II"
CDS
9031..9978
/gene="cgta-II"
/note="orf5-II; phase-variable gene due to homopolymeric
G-tract"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAL06007.1"
/db_xref="GI:15718507"
/translation="MKKILSIYKRYISISKLVLDNEHFIKENKNIYKHKHGGFFDPD
EKADKXPLNPPWGFYRVKNEALTURVLESILPALQGIITAYNDCDDGSELIIEFC
KQYFNFAKPYKVDNPNKNEKNLYSYNWAASFPLDEWFKIDVDHYDDAKL
YKSYRIDQENKALCYPRINFIILNGNIYVQNSNGYFISGGDQLIKRRNSSFIER
VSKSDMDPKGLIEELYSEQOVLQSOVKVILQAPLLOWHFPALKVRRNDYQOYLILS
LEEQAFHRKSKKAKIDFAMLKRPVIEQLKKQGEIK"
9975..10640
/gene="neua1"
CDS
9975..10640
/gene="neua1"
/codon_start=1
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAL06008.1"
/db_xref="GI:15718508"
/translation="MSLAIIPARGSGKIGKNKLVLLNNKPLIYVYTIKAALNAKSISK
```

Query Match 92.0%; Score 805.6; DB 15; Length 12388;  
Best Local Similarity 95.0%; Pred. No. 8e-212;

	Matches	832;	Conservative	0;	Mismatches	44;	Indels	0;	Gaps	0;
QY	1	ATGAAAAAGT	TATTATTCTG	TAAATGGA	CGAACGTT	TAAAGAAAT	TGATTATT	CAAGG	60	
Db	6001	ATGAAAAAGT	TATTATTCT	GGAATGGA	CGAACGTT	TAAAGAAAT	TGATTATT	CAAGA	6060	
QY	61	CTACCAAAAT	TGATTTT	TGATGTA	TGATGTA	TGATGTA	TGATGTA	TGATGTA	120	
Db	6061	CTACCAAAAT	TGATTTT	TGATGTA	TGATGTA	TGATGTA	TGATGTA	TGATGTA	6120	
QY	121	CTTGTAAAAA	TGCAAGCAG	TGTTT	TACACCC	TAAATTT	CTTCTTT	TGACCAAT	180	
Db	6121	CTTGTAAAAA	TGCAAGCAG	TGTTT	TACACCC	TAAATTT	CTTCTTT	TGACCAAT	6180	
QY	181	ACTTTAAAAA	CATTTAAT	CCAAAAT	CAAGAAT	TAGAGAC	CGCAAT	TAAATTT	240	
Db	6181	ACTTTAAAAA	CATTTAAT	CCAAAAT	CAAGAAT	TAGAGAC	CGCAAT	TAAATTT	6240	
QY	241	TACAACCAAG	CTCATCTA	GAAAAAT	TGAAAAA	CTTTT	TGAAAAA	CTTTT	300	
Db	6241	TACAACCAAG	CTCATCTA	GAAAAAT	TGAAAAA	CTTTT	TGAAAAA	CTTTT	6300	
QY	301	GATGCTCAT	TGCGGAT	GATGTTT	TAAACAA	CTTAAAGAA	TGCTTAT	TGCTTAT	360	
Db	6301	GATGCTCAT	TGCGGAT	GATGTTT	TAAACAA	CTTAAAGAA	TGCTTAT	TGCTTAT	6360	
QY	361	TTTTCAGGAA	TTTATTT	CAATCAAG	AAATTT	TACCTCAG	GGGTCT	TATATG	420	
Db	6361	TTTTCAGGAA	TTTATTT	CAATCAAG	AAATTT	TACCTCAG	GGGTCT	TATATG	6420	
QY	421	ATAGCCCTAG	GATACAAA	GAAATTT	TATCTT	CGGGAAT	TGATTTT	TATCAAAAT	480	
Db	6421	ATAGCCCTAG	GATACAAA	GAAATTT	TATCTT	CGGGAAT	TGATTTT	TATCAAAAT	6480	
QY	481	TCTTATGCT	TTTGTAT	ACCAAG	AAATTT	CTTTT	TAAAC	TAGCCCT	540	
Db	6481	GGGTATGCT	TTTGTAT	ACCAAG	AAATTT	CTTTT	TAAAC	TAGCCCT	6540	
QY	541	GATCGCTCG	CACTATAT	CGGACAT	AGTATA	AAAAAT	TACAGAT	TAAAGCT	600	
Db	6541	GATAATTC	ACATATAT	CGGACAT	AGTATA	AAAAAT	TACAGAT	TAAAGCT	6600	
QY	601	GAAGAAAA	CTTACAAA	ATAAAA	CTTAT	TGCTTAT	GTCTAAT	AGTCTTT	660	
Db	6601	GAAGAAAA	CTTACAAA	ATAAAA	CTTAT	TGCTTAT	GTCTAAT	AGTCTTT	6660	
QY	661	ATAGAACTAG	CGCCAAAT	TTAAAT	TCAAAAT	TTTAT	CATACA	AGAAAAA	720	
Db	6661	ATAGAACTAG	CGCCAAAT	TTAAAT	TCAAAAT	TTTAT	CATACA	AGAAAAA	6720	
QY	721	AAAGATATAT	CTCATCT	TCTTAG	TGAGGCT	TATGGA	AAAAAT	TTTCAAAAAA	780	
Db	6721	AAAGATATAT	CTCATCT	TCTTAG	TGAGGCT	TATGGA	AAAAAT	TTTCAAAAAA	6780	
QY	781	AAAAAAAT	AAAAAT	TAAAGAA	ATGTTT	TATCAAG	TTGATA	TAAAGAT	840	
Db	6781	AAAAAAAT	AAAAAT	TAAAGAA	ATGTTT	TATCAAG	TTGATA	TAAAGAT	6840	
QY	841	CCTAGTGAT	ATATAAG	CATTATT	TTTCAAG	GAAAAATAA	876			
Db	6841	CCTAGTGAT	ATATAAG	CATTATT	TTTCAAG	GAAAAATAA	6876			

RESULT 67  
AY854153  
LOCUS  
DEFINITION  
Campylobacter jejuni strain GB5 lipooligosaccharide biosynthesis  
locus, partial sequence.  
ACCESSION  
AY854153  
VERSION  
AY854153.1  
GI:61387361  
KEYWORDS  
Campylobacter jejuni  
ORGANISM  
Campylobacter jejuni

```

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae, Campylobacter.
REFERENCE
1 (bases 1 to 12403)
AUTHORS Gilbert,M., Brochu,D. and Karwaski,M.-F.
TITLE Sequencing of the Campylobacter jejuni GB5 LOS locus
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 12403)
AUTHORS Gilbert,M., Masotti,M. and Karwaski,M.-F.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-2004) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
source
1..12403
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="GB5"
/serotype="HS:4/64"
/db_xref="taxon:197"
<!--12403
/notes="lipooligosaccharide biosynthesis locus"
-->
CDS
<1..353
/codon_start=3
/transl_table=11
/product="putative heptosyltransferase I"
/protein_id="AAK45333.1"
/db_xref="GI:61387362"
/translation="EFAKEVNLNSGIDETHIELAPKFSLEELMAFTKQMDLIIGNDSG
EDLPAFNKASITIFGATPSYRNAFQTHINKIIDAGKIQNAKHIDKSDFCITRIE
EPIFKLAGLNEK"
343..1230
/codon_start=1
/transl_table=11
/product="putative lipid A biosynthesis acyltransferase"
/protein_id="AAK45334.1"
/db_xref="GI:61387363"
/translation="MKNSDRIYLSLYILKEFVTFMPECTLHFLALIVARTFYLNKK
.HRKIIDNLQICFPYQTKERDKLSKIENFAQFGIDCQNTTKREILNKVNFIN
ENLIDALAKRPIIFTTHYGNWEILSLAYAAKGAIAIVGKKLSEIMYELISQSR
TFQDIELIDKGGIRQMSALKKRGTLGILTDQCAENESVRLKFFKNEVNYQMSAL
IAQSNALIIIPVAYKEDGKPCIEFFKAKDSQNASLEELTYLQAQSCBEMIKKRPWEY
FEHRRFASYNKEIYKDAK"
1227..2774
/codon_start=1
/transl_table=11
/product="putative two-domain glucosyltransferase"
/protein_id="AAK45335.1"
/db_xref="GI:61387364"
/translation="MNLKQISVIIIVKNAQTLLLECLNSLKDFDEIILLNNESSDNTL
NIANFKKDFANLHIYHSAFIGFALKNLALSAYAKNDWILSIDADEVLENECIEELKN
LEQEDNIIALSRLNLYKGWIKACGWPDYVWRIFNKFTFRDNLVLHESLVLPNSV
KTIHLKGLKHAFRDISHLINKQHYSSLWAKQNIHKSGVLKANLRAFMTFRNYF
LNGFLYXGKFIISVNCALGTFFKYLKLEQKPKTCALIIITYNOKERLKLVLDD
SVKNLTPLNEVLJADGSKEDTARLJEYQKDFCPLKHIWQDEGFRLSKSNKTI
KADSEYIIIVDGMILEKDPKIEHFEPAQRRLFLQSRVILNEKESSEILNKDDYRI
AFNKQDFKSNKSLAKMLYLSLKKDEKIFPKTTLIKIGCNCSFPFTDKLDGFN
ENFTGWGREDSSEFVARELFNKGIFRRLKFKAIYAHYIHEENSKMLESNHQIYLDTIK
NKTSWR"
2778..3971
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAK45336.1"
/db_xref="GI:61387365"
/translation="MKKIGVVIPIVNVKYLRECLDSVINQTYNLTMLVNDGSTDE
NSLNAIKTYLKDRIITLFDKNGKLSARNVGLIEYSGKYLKKNKQTHIKENSLIEF
QLDNNPNYIKAYKSOAFNEKDLNTFTYPSIDYIIFLSDNFWELNLEPCVIRM
KNVDLWFDHQTQIEDNIKKNKKTREIEIFNKEECIITPKYANRALSVGRSDISFG
WNGMIDFNFLQIKLKFNLINEDIHFGIILFASANKIYVLSQLYLCIRLRANSISN
HDKKITRANVSEYFKDIETFGENAKAKNVLKAASRIITALKLIEFPKDKSENALA
IKETFLPCYAKALMIKKFKDPLNLKEQLVLKPFITQTKLPYDIWKEFWQKIKFIKN
NKIKN"
4020..5078

```

---

```

/codon_start=1
/transl_table=11
/product="putative
beta-1,4-N-acetylglactosaminyltransferase"
/protein_id="AAK45337.1"
/db_xref="GI:61387366"
/translation="MLFQSYFVKIICLPIPRKIRHKIKTKTFLLNKIQDKIDSYLPK
KTLAQINKYNNEDLIKLNAIIGBEHGYFNVDEKSDPKPLNFWAFIRKNEAITL
KASLESLPAIORGVIGVNDCTDGESEIILRECKQYPSFIPKYPVEIQONPKSEEN
KLYSYNVASFIKDEWLIKIDVDHYDAKLKYSFYIPKHYHVISRVIDFLFNE
EKFYVNGEIGILKAPDCLAIQNTNLFWEILIEDDTFKWNTAKNLENASKEIIL
KVRNRIYFTTELNNHFFPIKRYKNDYKQLNWNVSLDDFIKYNKEKLNQIDFKMLEY
KTLKVKYKLTSTSDKI"
complement (5730..5942)
/notes="inactive protein"
/codon_start=1
/transl_table=11
/product="truncated beta-1,3-galactosyltransferase"
/protein_id="AAK45338.1"
/db_xref="GI:61387367"
/translation="MPKISIIILPTFNVKVIATAKALESCINQSFKDIEIIVDDCGSDK
SIDIAKEYAKKDRIKIIHNEENLF"
6031..6822
/codon_start=1
/transl_table=11
/product="putative alpha-2,3/8-sialyltransferase"
/protein_id="AAK45339.1"
/db_xref="GI:61387368"
/translation="MKKVIIAGNGPSLKEIDYSRLPNDFDVRCNQFYPEDKYLGKK
CKAYVNPSPFFEQVYTLKLIQNOEYETELIVCSFNPLTHIESENFKNFYDFPDA
HLGYDFKQKLEFNAYPKFHEIYNQRTSGIYMCALVAILGYKKIYLSGIDIFYQNGS
SYAEDTKOKNLLKLVSNFKNDSHYIGHKSTDLKALFELEKTYIKLYCLCPNSLLA
NFIELAPNLNSNFIKKKKTITLKLYSILLVRLMENFQKILPLKK"
6741
/product="results in truncated putative alpha-2,3/8
sialyltransferase"
/replace=""
6742..6743
/product="results in extended C-terminus of putative
alpha-2,3/8 sialyltransferase"
/replace="a"
6909..7949
/codon_start=1
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAK45340.1"
/db_xref="GI:61387369"
/translation="MKEIKIONIIIEEKAPLVPEIGINHNGSLKELAKIMVNAFSA
GAKTIKHOITHIVEDMSKAAKVIIPGNAKISIVEIMOKCALDYKDELALKEYTEKLGL
VYLSFPFRAGANRLDMGVSAFKISGECNNYPLIKHIAAPKPMIVSTGWNISIESI
KPTVKILLDNEIPFVLMHTNDIYTPHNLVRLNLMLELKKFSCWVGLSDHTDNLAC
LGVALGACVLERHFTDSMRSGDPDIVCSMDTQALKELIIQSEQMAIMRGNNESKAA
KQGVATIDFAPASVVSIIKIDIKKEVLSMDNIWVRPGLGGISASEFENILGKKALRDI
ENDTQLSYEDFA"
7946..9070
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine-6-phosphate
2-epimerase"
/protein_id="AAK45341.1"
/db_xref="GI:61387370"
/translation="MKKILFITGTRADYSKISLMYRVONSSEFELYIATGMLSKN
FGYTVYKNGFNKNYIEFYNDKYQTDKALATTIDGSRVNEUKPDLIIVHGDRIG
FLAAAIYVGNALNLIHVAIEGIEISGTTDDSLRHAIKSLAHILHVNDEFKRLMOGI
EDEKSIIFIIGSPDELLNDNKISLNEAKKYDINVENTALLMFHPVTYTTISKMOAG
NLVAKILOSNNYIVIPYNDIGLFFELILOSVEELKNRPLKPPSRFFYFTLLKNA
DFIGNSSCILKEALYLTAGTILVSRQNGRLNENTLVNANSDELKAINTIHKKQ
DIFSAKLIELOSSKULFFELYQSGEFFKLSQTKVFKDIKHA"
9063..10010
/codon_start=1
/transl_table=11
/product="putative
beta-1,4-N-acetylglactosaminyltransferase"

```





misc_feature	LEEFQAFHRKSKEAKKIDFAMLRKRPVIEIQILKKFQGEIK"	
	9633..9642	
	/note="homopolymeric G-tract"	
	9633..9642	
variation	/note="G(10)-G(11) polymorphism; G(11) results in truncated putative beta-1,4-N-acetylglactosaminyltransferase"	
	/replace="G(11)"	
	9980..10645	
	/codon_start=1	
CDS	/transl_table=11	
	/product="CMP-Neu5Ac synthetase"	
	/protein_id="AAR82861.1"	
	/db_xref="GI:40217891"	
CDS	translation="MSLAIIPARGSGIKKNKLVLLNNKPLIVYTIKAALNAKSISK VLSSDSVILNVAQSNVDILKRPISLAQDDTTSDKVLHAKFKYKYEDVIFLOPT SPURNTDKAFNLKNSNANALISVSECNKLKAFVNDYDGLAGICNDVEYFMP ROKLPKTYMSNGAIYILKIKELFNPNPSFLQNKTRYFLMDESSSLDIDCLEDLKKVEQI WKK"	
	10633..11463	
	/note="translation dependent on homopolymeric G-tract polymorphism"	
Query Match	89.0%; Score 779.8; DB 15; Length 12370;	
	Best Local Similarity 93.7%; Pred. No. 1e-204;	
	Matches 824; Conservative 0; Mismatches 52; Indels 3; Gaps 1;	
Qy	1 ATGAAAAAGTTATTTCGTGGAAATGGACACAGTTTAAAGAAATGATTATTCAAGG 60	
	Db 6003 ATGAAAAAGTTATTTCGTGGAAATGGGCAAGCTTAAAGAAATGATTATTCAAGG 6062	
Qy	61 CTACCAATGTTTTGATGTTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120	
	Db 6063 CTACCAATGATTGATGTTATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 6122	
Qy	121 CTTGTAAAAATGCAAGACAGTGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180	
	Db 6123 CTTGTAAAAATGCAAGACAGTATTTTACAACTCTAGTCTTTTTTTTGAACAAATACTAC 6182	
Qy	181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAACTAATTTATGTGTTCTAAT 240	
	Db 6183 ACTTTAAACATTTAATCCAAATCAAGATATGAGACAGAAATTAATCGTGTGTTCTAAT 6242	
Qy	241 TACAACAGCTCATCTAGAAAATGAAAATTTGTAAAACTTTTACGATTAATTTTCT 300	
	Db 6243 TTTAATCTAATCTATAGAAAAGCGAAATTTTAAAGAACTTTTATGATTAATTTTCT 6302	
Qy	301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360	
	Db 6303 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 6362	
Qy	361 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTCAGTAGCC 420	
	Db 6363 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGAATCTATATGTGTCAGTAGCC 6422	
Qy	421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGGTCA 480	
	Db 6423 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGATCA 6482	
Qy	481 TCTTATGCTTTTGATACCAACACAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540	
	Db 6483 TCTTATGCTTTTGATACCAACAAATAATCTTTTAAATGGTTCTTAAATTTTAAAT 6542	
Qy	541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTAAATTTCTA 600	
	Db 6543 GATTAATTCACACTATATCGGACACAGCAAAATACAGACCTTAAAGCTTTTAAATTTCTA 6602	
Qy	601 GAAAAAATTCACAAAAATAAATATATGCTTATGTCCTAATAGTCTTTTACGAAATTTT 660	
	Db 6603 GAAAAAATTCACAAAAATAAATATATGCTTATGTCCTAATAGTCTTTTACGAAATTTT 6662	
Qy	661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCAT--ACAAGAAAAAATAACTAC 717	
	Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCAT--ACAAGAAAAAATAACTAC 717	
Db	6663 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATATAAAAAAATAAATAACTAC 6722	
	Qy 718 ACTAAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAT 777	
Db	6723 ACTAAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAT 6782	
	Qy 778 TTTAAAAAATAAATAATTAAGAAAAATGTTTATTATCAAGTTGATAAAGATCTTATTAAGA 837	
Db	6783 TTTAAAAAATAAATAATTAAGAAAAATTTTATTATCAAGTTGATAAAGATCTTATTAAGA 6842	
Qy	838 TTACTAGTATATATAAGCATTTATTTCAAAGGAAAAATAA 876	
	Db 6843 TTACTAGTATATATAAGCATTTATTTCAAAGGAAAAATAA 6881	
RESULT 69	BD134499 1293 bp DNA linear PAT 18-SEP-2002	
	LOCUS Lipopolysaccharide alpha-2,3 sialyltransferase of Campylobacter	
	DEFINITION jejunii and its uses.	
	BD134499 GI:23229444	
ACCESSION	JP 2002507424-A/1.	
	VERSION Campylobacter jejuni	
	KEYWORDS Campylobacter jejuni	
	SOURCE Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.	
ORGANISM	1 (bases 1 to 1293)	
	REFERENCE Gilbert, M. and Wakarchuk, W.W.	
	AUTHORS Lipopolysaccharide alpha-2,3 sialyltransferase of Campylobacter	
	TITLE jejunii and its uses	
JOURNAL	Patent: JP 2002507424-A 1 12-MAR-2002;	
	NATIONAL RESEARCH COUNCIL OF CANADA	
	OS Campylobacter jejuni	
	COMMENT PN JP 2002507424-A/1	
COMMENT	PD 12-MAR-2002	
	PF 22-MAR-1999 JP 2000538012	
	PR 20-MAR-1998 US 60/078891, 18-MAR-1999 US	
	PC MICHEL GILBERT, WAKREN W WAKARCHUK	
FEATURES	C12N15/09, C12N1/21, C12N5/10, C12N9/10, C12P19/26, C12N15/00, C12N5/ PC 00	
	CC Campylobacter jejuni OH4384 est-I gene alpha2,3- CC	
	sialyltransferase	
	PH Key Location/Qualifiers	
source	FT CDS (1)..(1293).	
	Location/Qualifiers	
	1..1293	
	/organism="Campylobacter jejuni"	
ORIGIN	/mol_type="genomic DNA"	
	/db_xref="taxon:197"	
Query Match	35.6%; Score 311.6; DB 2; Length 1293;	
	Best Local Similarity 61.1%; Pred. No. 2.7e-75;	
	Matches 522; Conservative 0; Mismatches 329; Indels 3; Gaps 1;	
Qy	1 ATGAAAAAGTTATTTCGTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60	
	Db 46 ATGCAAAATATAATCATAGCAGAAATGCCTAGCTTAAATAATTAATATATAAAGA 105	
Qy	61 CTACCAATGATTTTGAATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120	
	Db 106 CTGCTAGAGAAATATGATGTTTTTAGGTGTAACCAAGTTTATTTTGAAGATAAGTATTAT 165	
Qy	121 CTTGTGTAATAATGCAAGCAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180	
	Db 166 TTGGAATAAAGATTAAGCAGTATTTTTTAATCCTGGTGTCTTTTTTACACAGTATCAC 225	
Qy	181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240	
	Db 226 ACTGCAACAACTTATATCTATACTAAAAATAGATGAATAAAAAATATTTTGTCTCTACA 285	



```
Qy 241 TACAAACCAAGCTCATCTAGAAATGAAAATTTTCTGTAAGAACTTTTACGATATTTTCCT 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
286 TTTAAATTTACCTTTTATGAAAGCAATGATTTTATACATCAATTTTATAATTTTCCCC 345
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAA 360
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 GATGCAAACTTGGCTATGATTTTAAAGCACTTAAAGAAATTTATGCTTTATATAA 405
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 420
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 TACAAATGAAATTTATTTCAATTAAGAAATTTACTTGGCGCTCTATATGTTGCAATTTGCT 465
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 ATAGCCCTAGGATACAAAGAAATTTACTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
466 ATTGCATTAGGATATAAAACCATCTATTTATGTCGCAATGATTTTATGAAGGAGATGTT 525
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 540
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
526 ATTATCTTTTGAAGCTATGAGTACAAATATAAAACAACTTTCTCGGATATAA 582
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 GATGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTGAATTTCTA 600
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
583 GATTTCAAACTTGGCAATTTGCAATTTCAAGGAATACGATATAAGCAATTTGTTA 642
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 GAAAAAATTTACAAATAAACTATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
643 AAATCAATATACAAAGTTAATATCTACGCAATGTTGTAATGATTTCTTTTGGCAATCAT 702
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAATAAACTACACT 720
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
703 TTTCTTATCAATTAATTAATTAACAAATTTCAATTTAGAAATTAAGCAATTAATTTCT 762
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 AAAGATATACATACCTTTCTAGGCTTTATGGAATAATTTTCAAAATAATTTTAAATTTT 780
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
763 ATAAATGATATTTTATGACTGATAATCTCTCGCGTAAGTTTATATAAAATCAACTT 822
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
781 AAAAAATAAAATTAAGAAATTTTATTTACAGTTGATATAAGATCTATTAGATTA 840
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
823 AAAGCTGATAATAAAATTTATGCTTTAAATTTTATAATATTTCTTCTTAAAGATAATTTA 882
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
841 CCTAGTGATATAA 854
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
883 ATTAAATTTTAAA 896

RESULT 71
AR473907 LOCUS AR473907 1293 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 6689604.
ACCESSION AR473907
VERSION AR473907.1 GI:42712446
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Lipopolysaccharide .alpha.-2,3 sialyltransferase of Campylobacter jejuni and its uses
JOURNAL Patent: US 6689604-A 1 10-FEB-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1. 1293
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 35.6%; Score 311.6; DB 2; Length 1293;
Best Local Similarity 61.1%; Pred No. 2.7e-75;
Matches 522; Conservative 0; Mismatches 329; Indels 3; Gaps 1;
Qy 1 ATGAAAAAGTATTATTGCTGGAAATGAGCAACCAAGTTTAAAGAAATTTGATTTCAAGG 60
```

```
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
46 ATGCAATAATAATCATAGCAGGAAATGAGCTAGCCTAAATAATTTAATTATAAAGA 105
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 CTACAAATGATTTTGTATTTAGATGTAATCAATTTATTTTGAAGATAAAACTAT 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 CTGCTAGAGAAATATGATGTTTTTATAGGTGTAACCAAGTTTATTTTGAAGATAGTATT 165
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 CTTGGTAAAAATGCARAGCAGTGTTTTACACCCCTAAATTTCTTCTTTCAGCAATACTAC 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
166 TTAGAAAAAAGATTAAAGCAGTATTTTATCTCTGGTGTCTTTTACACAGTATCAC 225
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACCTAATTTATGTTCTTAAT 240
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
226 ACTGCAAACTATATATCTAAATAGATATGAAATAAAATAATTTTTCCTCTACA 285
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 TACAACTAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTTTTCCT 300
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
286 TTTAATTTACCTTTTATTTGAAAGCAATGATTTTTCATCAATTTTATAATTTTTCCTCC 345
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTATTATAA 360
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
346 GATGCAAACTTGGCTATGAGTTAATGAAACCTTTAAAGAAATTTTATGCTTATAA 405
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 420
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
406 TACATGAAATTTTATTTCAATAAAGAAATTTCTCGGGCTCTATATGTTGCAATTTGCT 465
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
466 ATTGCATTAGGATATAAAACCATCTATTTATGTCGCAATGATTTTATGAAGGAGATGTT 525
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 540
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
526 ATTATCTTTTGAAGCTATGAGTACAAATATAAAACAACTTTCTCGGATATAA 582
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 GATGCTCGCACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTGAATTTCTA 600
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
583 GATTTCAAACTTGGCAATTTGCAATTTCAAGGAATACGATATAAGCAATTTGTTA 642
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 GAAAAAATTTACAAATAAACTATATGCTTATGCTTATGCTTAAATAGTCTTTTACGAAATTTT 660
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
643 AAATCAATATACAAAGTTAATATCTACGCAATGTTGTAATGATTTCTTTTGGCAATCAT 702
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
661 ATAGAACTAGCGCCAAATTTAAATTTTCAATTTTATCATACAAAGAAATAAACTACACT 720
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
703 TTTCTTATCAATTAATTAATTAACAAATTTCAATTTAGAAATTAAGCAATTAATTTCT 762
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 AAAGATATACATACCTTTCTAGGCTTTATGGAATAATTTTCAAAATAATTTTAAATTTT 780
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
763 ATAAATGATATTTTATGACTGATAATCTCTCGCGTAAGTTTATATAAAATCAACTT 822
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
781 AAAAAATAAAATTAAGAAATTTTATTTACAGTTGATATAAGATCTATTAGATTA 840
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
823 AAAGCTGATAATAAAATTTATGCTTTAAATTTTATAATATTTCTTCTTAAAGATAATTTA 882
Qy ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
841 CCTAGTGATATAA 854
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
883 ATTAAATTTTAAA 896

RESULT 71
AR489130 LOCUS AR489130 1293 bp DNA linear PAT 15-MAY-2004
DEFINITION Sequence 1 from patent US 6709834.
ACCESSION AR489130
VERSION AR489130.1 GI:47256032
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Lipopolysaccharide .alpha.-2,3 sialyltransferase of campylobacter
```



[illegible]

DNA-directed RNA polymerase (EC 2.7.7.6) beta'-2 chain - Plasmodium falciparum  
 S72284  
 N:Alternate names: RNA polymerase rpoC2  
 C:Species: plasmodium falciparum  
 C:Date: 14-Apr-1998 #sequence\_revision 24-Apr-1998 #text\_change 09-Jul-2004  
 C:Accession: S72284  
 R:Wilson, R.J.M.; Denny, P.W.; Preiser, P.R.; Rangachari, K.; Roberts, K.; Roy  
 J. Mol. Biol. 261, 155-172, 1996  
 A:Title: Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*  
 A:Reference number: S72277; MUD:96346169; PMID:8757284  
 A:Accession: S72284  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-960 <WtL>  
 A:Cross-references: UNIPROT:Q25802; UNIPARC:UPI0000075E46; EMBL:X35275; NID:g1  
 A:Note: biosynthesis of this protein involves a -1 frameshift in the codon for  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January  
 C:Genetics:  
 A:Gene: rpoC2  
 A:Genome: plastid  
 A:Note: this apparently degenerate plastid is referred to as the apicoplast  
 C:Superfamily: DNA-directed RNA polymerase beta chain  
 C:Keywords: nucleotidyltransferase; plastid; transcription

## plast

[illegible]

D	b	604	WYKYLELLNNKYLVIIYYNNYIKDYLYKYNINLNLYFI--KNLFYN-----NNNFIFHNH
Q	y	233	IQEKNN-----YTKDILIPSEAYGKF-----SKNI 258
D	b	654	IYKNNYIYVNNNNLYQYNKGNILINNLLYNNKLFYNYINNVIYLNLDITIGLSINI 713
Q	y	259	NFKIKIKENV-----YVKLIKDLRLPS-----DIKH 287
			:   :   :



QY 217 -----LANFIELAPLNSNFIQ--EKNNYTKDILIPSSAYGKFSKNINFKKIKIKKE 267  
Db 1027 YKTFOWISELIFYSKNL--NYKIQNKIEKNYC-----YKNKISYKKKIKT 1071  
QY 268 -----NVY--YKLIKDLRLPSDIKHFKGK 291  
Db 1072 VNFFKNNLFQNNNSWFFLEWEYNTYILQIOETFTQITDVLEFYKKK 1122

RESULT 8  
E65140  
hypothetical 44.2K protein (gntR-ggt region) - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
A:Accession: E65140  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: E65140  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-392 <BLAT>  
A:Cross-references: UNIPROT:P46955; UNIPARC:UPI0000047CA0; GB:AB000420; GB:U00096; NID:97426617  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yhhZ  
C:Superfamily: Escherichia coli hypothetical 44.2K protein (gntR-ggt region)

Query Match 7.3%; Score 114; DB 2; Length 392;  
Best Local Similarity 22.4%; Pred. No. 0.67;  
Matches 69; Conservative 39; Mismatches 80; Indels 120; Gaps 16;

QY 25 FDFVRCNQFYEDKYY-----LGKCKKTVFTYTNFFFEQYTYLKH 64  
Db 87 FDFYKINFRGLKYYIYQLGAFLSAIHQIENQOLDTETITSY-----EPILCOH 139  
QY 65 LIQOYEYETELIMCSNYNOAHLEN-----ENFKY 93  
Db 140 LIANTEF-SYALPENYRNLEFLPNSKQNTNRFKTLNSKATGRLLAAGVYNGIEGRAD 198  
QY 94 TFYDYFPPDAHLGYDFFKOLKEFNAYKPHFIYFNQRIITSGVYMCVAIAIAGYKEIYLSGI 153  
Db 199 TAELKGGDAIKGYD--QI-----LNEK-TAGIAIATASILLTKR-----SNV 237  
QY 154 DFYQNGSSYADTKQENLLKLAPDFNDRSHYGHKNTDIKALEFLEKTYKIKLYCLCP 213  
Db 238 DTYTEINSY-----LGKLR-----GQCKLLD--GIDIIE-----IYIKRP 271  
QY 214 NSLLANF-IELAPLNSNFIQEKNNYTKDILIPSSAYGKFSKN--INFKK--IKIKEN 268  
Db 272 SKDLANLKEFNKTVRKFLIK-----LAKTSEASGRFNAEDLLMRKGNVPLWYN 322  
QY 269 VYKLIKD 276  
Db 323 VHKLSLD 330

RESULT 9  
T28423  
ORF MSV261 leucine rich repeat gene family protein - Melanoplus sanguinipes entomopoxvirus  
C:Species: Melanoplus sanguinipes entomopoxvirus  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
A:Accession: T28423  
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.  
J. Virol. 73, 533-552, 1999  
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
A:Reference number: Z20484; MUID:99102612; PMID:9847359  
A:Accession: T28423  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-624 <AFO>

A:Cross-references: UNIPROT:Q9YV17; UNIPARC:UPI000000EFC79; EMBL:AF063866; NID:g4049647;  
C:Genetics:  
A:Note: MSV261

Query Match 7.3%; Score 114; DB 2; Length 624;  
Best Local Similarity 22.3%; Pred. No. 1.2;  
Matches 69; Conservative 37; Mismatches 118; Indels 86; Gaps 15;

QY 9 NGPSLKEIDYSR-----LPNDFVFRCNQFYEDKYYLGKCKKTVFTYTNFFFEQY 60  
Db 343 NLDNLKELDISRNFNINISLPSKLITLINSYIKH-----TNFINDL 387  
QY 61 TLKHLIQOYEYTEL-----IMCSNYNOAHLENFVKTYFYDYPDAHLGYDFFKOLKE 114  
Db 388 NLKELDISRNFNINISLPSKLILKINISCDIKNTWFINDL-----DNLKELDI 436  
QY 115 FNAYKFKHEIYNQRIITSGVYMCVAIAIAGYKEIYLSGDIYFQNGSSYADTKQENLLKL 174  
Db 437 SNY-----NINIS--LPKTLTTLNISCNKNTNFIIN-----LDNLKE--LDI 478  
QY 175 APDFK-----NDRGHYIGHSKNT-----DIKALEFLEKTYKIKLYCLCPNSLL 217  
Db 479 SKNFNINISLPSKLITLINSYINFTNFINDLNKLKELD-LSNNYNINISL-PNTLI 536  
QY 218 ANFIELAPLNSNFIQEKNNYTKDILIPSSAYGKFSKNI--NFKKIKIKENV---Y 270  
Db 537 ELNISCYIKNTNFIINLDN--LKLEISKSLYDMILHNIFTNFKLKLKDKNTSFTFY 594  
QY 271 YKLIKDLRL 280  
Db 595 YKIKSKFKVI 604

RESULT 10  
T28418  
ORF MSV257 leucine rich repeat gene family protein - Melanoplus sanguinipes entomopoxvirus  
C:Species: Melanoplus sanguinipes entomopoxvirus  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
A:Accession: T28418  
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.  
J. Virol. 73, 533-552, 1999  
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
A:Reference number: Z20484; MUID:99102612; PMID:9847359  
A:Accession: T28418  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-707 <AFO>  
A:Cross-references: UNIPROT:Q9YV15; UNIPARC:UPI00000F54AF; EMBL:AF063866; NID:g4049647;  
C:Genetics:  
A:Note: MSV257

Query Match 7.3%; Score 114; DB 2; Length 707;  
Best Local Similarity 19.1%; Pred. No. 1.3;  
Matches 98; Conservative 46; Mismatches 126; Indels 244; Gaps 22;

QY 6 IAGNGPSLKEIDYSRLPN--DFDVFRCN--QVFEEDKY----- 39  
Db 182 ISNNEQLSKKIDYLRPLNLDNFCNCSNKSFNFLNNFTKLQILDISNENINWYALPLP 241  
QY 40 -YLGK-KCKTVFTYTNFFFEQYTYLKLH-IQNO-----EYETELIMCSNYNOAHLENF 91  
Db 242 PFLIKVNCSCGNTNSDFLRYVDNLEELDISNPDLIKIEYMPPTKIKKLISECYIKDNF 301  
QY 92 VKTF-----YDYF----- 99  
Db 302 LKGLNQLNELNISYNPNYFRKSIETIDNLSNTLIKLYLRKCNKNVSNFKNLINLQEL 361  
QY 100 -----PD-----AHGYDFFKOLKEFN-----AYFKFH 122  
Db 362 VISENRQVNNLNPDELISLNSCKKINIEFPKLIKELNLSKTYISNKNLPKSLIKLD 421  
QY 123 EIYFNOR-----ITSGVYMCVA-----IALGYKEIYLSGIDFYQN--- 158

Db 422 ISVCNLKNDCLNKLNLKELDISGNWEICFSSLPKSIKIKCKNCNINTWDFLNLNDL 481  
QY 159 -----GSSYADPTQENLLKAPDPKDRSHYIGHSKNTDIALKFLBK--- 202  
Db 482 QBLDMSVNYIININKSRRIDNLPKLIK---NINNCDCAYISPKGLDNLEELSMIRNTCL 539  
QY 203 -----TYKI-----KLYCLC-----PNSL----- 216  
Db 540 IDYLPKSLIKLHGHTFKIDKLNKLTTLCLPGIINKINTEIDPKSLFIELDISTCKYIINE 599  
QY 217 -----LANFIELAPLNSNFIQ-----EKNYTKDIL---IPSEAYG 252  
Db 600 DCLKNLYNLKELTMSINCNIKQTLPESLIKIKICKLCKYSYDDFKIDYDKKLNLSNKG 659  
QY 253 KP-----SKNINPKIKIKENVY-YKLKDL 277  
Db 660 KYISFLPEIKKNYNNAFKIHLIDNIYNLKKIKSI 693

## RESULT 11

D90129  
C:Species: hypothetical protein orf1613 [imported] - Guillardia theta nucleomorph  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
A:Accession: D90129  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; MUID:11323671; PMID:11323671  
A:Accession: D90129  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1613 <DOU>  
A:Cross-references: UNIPROT:Q98S55; UNIPARC:UPI00000A244C; GB:AF083031; NID:gl3794350; P  
C:Genetics:  
A:Gene: orf1613  
A:Map position: 3  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 7.3%; Score 114; DB 2; Length 1613;  
Best Local Similarity 22.4%; Pred. No. 3.5;  
Matches 73; Conservative 44; Mismatches 93; Indels 116; Gaps 18;

QY 46 KTVFVTPNFFEQYTL---KHLIQOEYETELIMCS--NYNQAHLENENF-----VKT 94  
Db 1313 KTLHTIIVSVNTPYFIKKCDKYMFMFKRIFLDSVNYNY-KISYEIFFQISWVTS 1371  
QY 95 FYDYFPDAHL-----GYDFPKQLKEFNAYF---KPH-----EY-F 126  
Db 1372 FSNLPFKTNLKMKNLFKFKFNLIILKELNFSATFLILKFLNLSINTELLKFYF 1431  
QY 127 NQRTSGVYMCVAIALGYKEIYLSGIDPYQNGSSYAFDTKQENLLKLAPDPKDRSHYI 186  
Db 1432 YOKPKNFYLI-----FSKNGSFI-----SLHLKFLPENKYQSKL 1470  
QY 187 GHSKNTDIK-----ALEFLE-----KTY--KIKLYCUC 212  
Db 1471 S-SKYFTTKFFFEILYFFLQSPFKFLNFFSNFNLNDNKKRLLYFTQYLLKLIYSEI 1529  
QY 213 P-NSLANFIELAPLNSNFI-IOEKNNTYKDILIPSSSEAYGKFSKNINFKIKIKENVY 270  
Db 1530 PTLNLKQINVLKMLHLEFISIQ-----ILKKFLNMLNKFNFYRAKDYNSYS 1580  
QY 271 YKLIK-----DLLRLPSDIKH 287  
Db 1581 FKITKRFIFNYSGLFNITKDIEF 1606

## RESULT 12

H82926  
conserved hypothetical UUI43 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: H82926  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a m  
A:Reference number: A82870  
A:Accession: H82926  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1272 <GLA>  
A:Cross-references: UNIPARC:UPI00000C1BB3; GB:AE002115; GB:AF222894; NID:G6899102; PID  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UUI43  
A:Genetic code: SGC3

Query Match 7.3%; Score 113.5; DB 2; Length 1272;  
Best Local Similarity 24.9%; Pred. No. 2.9;  
Matches 73; Conservative 33; Mismatches 88; Indels 99; Gaps 18;

QY 60 YTLKHLIQOEYETELI---MCSNYNQAHLE--NENFVKTFDY----- 98  
Db 373 FKIKHINIDSSNKAIIELOSSSDNTKLEANNKQLIKSYDYNPWSKIVSYEKKDN 432  
QY 99 ---FPDAHLGYDFPKQLKEFNAYFKFHEIYFNQRTSGVYMCVAIALGYKEIYLSGIDF 155  
Db 433 NKMIFDLH---DFPKQLKT---FIITHIRFDDNITS-----LG--KIKENSFEY 473  
QY 156 YONGSSYA-----FTKQENLLKLAP-DPKDRSHYIGHSKNTDIALKFLKLEYKIK 207  
Db 474 YQNDKEYLLKSLKYFDIKENQLYGSACFNEND-----DFKILK--NKTFFVK 520  
QY 208 LYCLCPNLSLANFIELAPLNSNFIIOEKN-----NY-----TKDILIP 246  
Db 521 YEIDTKNLIILNKYIPLNKYINVDK---KNLAQPKIVNVFGLNLYKLESIKIVNKNLLP 576  
QY 247 SSEAYG-KFSKNINFK---KIKIKENVYK-----LIKDLLRLPS-DIKHVKF 289  
Db 577 YNDHVNIOANNTNFSVWHKVYPKQNIINEFFDDSLIKDEFNLKIDFNHLWK 629

## RESULT 13

B90083  
gamma-tubulin [imported] - Guillardia theta nucleomorph  
C:Species: nucleomorph Guillardia theta  
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: B90083  
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re  
A:Title: The highly reduced genome of an enslaved algal nucleus.  
A:Reference number: A99082; MUID:11323671; PMID:11323671  
A:Accession: B90083  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-424 <DOU>  
A:Cross-references: UNIPROT:Q9SE4; UNIPARC:UPI0000091BF1; GB:AF165818; NID:gl3794424;  
C:Genetics:  
A:Gene: tubG  
A:Map position: 1  
A:Genome: nucleomorph  
C:Keywords: nucleomorph

Query Match 7.2%; Score 113; DB 2; Length 424;  
Best Local Similarity 20.3%; Pred. No. 0.87;  
Matches 68; Conservative 54; Mismatches 99; Indels 114; Gaps 15;

QY 39 YILGKKCKTVFYPNPFEEQ-----YYTLKHLIQN-QEYETELIMCSNYNQAHLENFV- 92  
Db 35 FSLGKNSTP---NDNVFFESNESFFTPRTIIFLSDRDFNYIMKSNYSKMYDKNRHFL 91  
QY 93 -----KTFVDYFPDAHLGYDFK-KQLKEFNA--YFK----- 120

Db 92 NMTGNSWLGYYEGISNCNVLDNLRKRIRKEMNSVKYFNVSINGGTGAGLSSYLLEY 151  
Qy 121 -----EHEIYENQRTSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFD 165  
Db 152 IRNYPKSFNCCSIFPDLYQNTQTFOPYSNVLSIA--WQGLYCDNSIFFQNHAI----- 205  
Qy 166 TKOENLLKLAPDFKNDRSHYIGHSKNTDIKALEFLEKTYKI-KLYCLCPNSLLANF----- 220  
Db 206 --ENLL-----TKNQINNELSRKVNYYIGKTIISIIMNTLNHNFTLET 246  
Qy 221 ----IELAPNLSNF-----IIQEKNNYTKOILIPSSSEAYGKFSKNI 258  
Db 247 LISPLIYNVPLNLFSTINIDILSKLRKRPSEMIENKTQNLNETSMNLDLKGNYLSSI 306  
Qy 259 -----NFKKIKIKENYVKLIKDLRLPSPDKHYF 288  
Db 307 ELSNRNFYKL-FANSIYDKIFDYQTRFYSQQLNQY 340

RESULT 14  
T28180  
hypothetical protein ORF19 - Melanoplus sanguinipes entomopoxvirus (strain Tuscon)  
C:Species: Melanoplus sanguinipes entomopoxvirus  
A:Variety: isolate Tuscon  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T28180  
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.  
J. Virol. 73, 533-552, 1999  
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
A:Reference number: 220484; MUID:99102612; PMID:9847359  
A:Accession: T28180  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-437 <AFO>  
A:Cross-references: UNIPROT:Q9YW73; UNIPARC:UPI00000F78FA; EMBL:AF063866; NID:g4049647;  
A:Experimental source: isolate Tuscon  
C:Genetics:  
A:Note: MSV019

Query Match 7.2%; Score 112.5; DB 2; Length 437;  
Best Local Similarity 19.5%; Pred. No. 0.98;  
Matches 55; Conservative 50; Mismatches 74; Indels 103; Gaps 14;  
Qy 16 IDYSRLP---NDFVFRNCQFYFEDKY-YLGKKCKTVETYNFRFFEQYVTLKHLIQNQ-- 69  
Db 231 LDISITPLKYNINNIETFLDNKYNVINDOIQLFHS-----FLDKHKIKKRYN 282  
Qy 70 -----EYETELIMCSNNOAHLENENFVKTFYDFPDHILGYDFFKQLKEFNAYFZH 122  
Db 283 RFIILFNKYKNELTQNSQNRRIYDNKILYSRIKYFNDENILYN-----YFEPK 332  
Qy 123 EIVFNQRTSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKOENLLKLAPDFKNDR 182  
Db 333 DLISNDKL-----IVLI-----NNIS-VPELKDVINIIVKP----- 362  
Qy 183 SHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKN-- 238  
Db 363 KKYIRKOLRKAIKTIIDAI-----INSN--LQAGNNDKEY 394  
Qy 239 -YTKDIL-----IPSSEAYG-KFSKNINFKKIKIKENVY 270  
Db 395 EYDENILHLIANEFNLSNNECENIKYNKDKYSYVKLYHNIY 436

RESULT 15  
E82912  
hypothetical protein U259 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: E82912  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A:Reference number: A82870  
A:Accession: E82912  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-621 <GLA>  
A:Cross-references: UNIPARC:UPI00000C1BEE; GB:AE002123; GB:AF222894; NID:g6899229; PIDN  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U259  
A:Genetic code: SGC3

Query Match 7.1%; Score 111.5; DB 2; Length 621;  
Best Local Similarity 22.3%; Pred. No. 1.7;  
Matches 67; Conservative 43; Mismatches 114; Indels 77; Gaps 14;  
Qy 16 IDYSRLPN---DFDVFRNCQFYFEDKYLGKKCKTVFY-----TPNFFPEQY 59  
Db 246 IDYHREDNFKDILDYKKN---YEDLYYDEKLNSSIYQQGLDKINKEYEYVNNFN--F 300  
Qy 60 YTLKHLIQNOYETELIMCS-----NYNQAHLENENFVKTFY 96  
Db 301 NTKKLPFQKFSYQMFLEKDKLGFKRYVLNRIDQKEVRARDHFNDEKKQKNKYFRSF- 359  
Qy 97 DYFPDAHILGYDFFKQLKEFNAYFKPHEIYFNQRTSGVMCAVAIALGYKEIYLSGIDFY 156  
Db 360 ----DKELNY--FQKYIDHNLIDININKISKEAVITQKIIKNLLLDAFNNNDVQITKYLQ 413  
Qy 157 QNGSSYAFDTKOENLLKLAPDFKNDRSHYIGHSK-----NTDIKALEFLEKTYKIKLYCL 211  
Db 414 QNDKEY-----QNKLIKQILIEFSN-----YFKNQKTPRDSNEYKQKLDYKIFSQNWY-- 462  
Qy 212 CPNSLLANFIELAPNLSNFIIOEK--NNYTKDILIPSSSEAYGKFSKNINFKKIKIKENVY 270  
Db 463 ----FILMNLKHFELSFNNWFTLPQKINPITKEMLGPE-----KYLKRL---KLLKPYEDDY 513  
Qy 271 Y 271  
Db 514 F 514

Search completed: July 19, 2006, 13:48:07  
Job time : 20.2031 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:30:11 ; Search time 137.396 Seconds  
(without alignments)  
1959.153 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIIAGNGPSLKEIDYSR.....KLIKOLLRLPSDIKHYFKGK 291

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 92501592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	291	2	Q93D05 CAMJE
2	1559	99.7	291	2	Q919Q5 CAMJE
3	1553	99.3	291	2	Q9F0M9 CAMJE
4	1521	97.3	291	2	Q93MQ0 CAMJE
5	1521	97.3	291	2	Q91AK3 CAMJE
6	1517	97.0	284	2	Q50FZ0 CAMJE
7	1507	96.4	291	2	Q938X6 CAMJE
8	1492	95.4	291	2	Q93CZ5 CAMJE
9	1451.5	92.8	292	2	Q5DT12 CAMJE
10	1188.5	76.0	263	2	Q2TKA8 CAMJE
11	804	54.4	294	2	Q7BP25 CAMJE
12	744.5	47.6	326	2	Q4QM36 HABI8
13	739	47.3	430	2	Q2RGF1 CAMJE
14	736.5	47.1	430	2	Q32VR8 CAMJE
15	730.5	46.7	320	2	Q4QNI8 HABI8
16	688.5	44.0	303	2	Q9CLP3 PASMU
17	485.5	31.0	231	1	I352 HAEIN
18	169	10.8	44	2	Q4HGT1 CAMCO
19	152.5	9.8	1014	2	Q8RIY7 CAMJE
20	152.5	9.8	1014	2	Q8RN21 CAMJE
21	152.5	9.8	1176	2	Q51CW9 ENTHI
22	145.5	9.3	673	2	Q4HEM3 CAMCO
23	135	8.6	849	2	Q815V7 PLAF7
24	134.5	8.6	2755	2	Q81HP8 PLAF7
25	133	8.5	271	2	Q3NR31 SHEFR
26	129.5	8.3	615	2	Q4XIN2 PLACH
27	129.5	8.3	919	2	Q4YZ88 PLABE
28	129	8.2	417	2	Q9ALTO CAMJE
29	129	8.2	567	2	Q4XYM5 PLACH
30	129	8.2	1639	2	Q4YVD8 PLABE
31	128.5	8.2	1347	2	Q4Z0S9 PLABE

32	128	8.2	682	2	Q4YZJ1 PLABE	Q4YZJ1 plasmodium
33	128	8.2	713	2	Q6BT89 DEBHA	Q6BT89 debaryomyce
34	128	8.2	819	2	Q6LF52 PLAF7	Q6LF52 plasmodium
35	128	8.2	842	2	Q815V4 PLAF7	Q815V4 plasmodium
36	128	8.2	2772	2	Q7RG21 PLAYO	Q7RG21 plasmodium
37	127.5	8.2	560	2	Q81JN4 PLAF7	Q81JN4 plasmodium
38	127.5	8.2	804	2	Q54TF7 DICDI	Q54TF7 dictyosteli
39	127	8.1	885	2	Q81EM9 PLAF7	Q81EM9 plasmodium
40	126	8.1	421	2	Q812T8 PLAF7	Q812T8 plasmodium
41	126	8.1	551	2	Q81JG5 PLAF7	Q81JG5 plasmodium
42	125.5	8.0	515	2	Q652W7 BORG	Q652W7 borrelia ga
43	125.5	8.0	601	2	Q891W3 CLOTE	Q891W3 clostridium
44	125.5	8.0	803	2	Q4YST0 PLABE	Q4YST0 plasmodium
45	125	8.0	2568	2	Q81AL8 PLAF7	Q81AL8 plasmodium

## ALIGNMENTS

### RESULT 1

Q93D05 CAMJE	PRELIMINARY; PRT; 291 AA.
ID Q93D05 CAMJE	
AC Q93D05	
DT 01-DEC-2001	integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001	sequence version 1.
DT 07-FEB-2006	entry version 13.
DE Alpha-2,3-sialyltransferase.	
GN Namescst-II;	
OS Campylobacter jejuni.	
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;	
OC Campylobacteraceae; Campylobacter.	
OX NCBI_TaxID=197;	
RN [1]	
RP NUCLEOTIDE SEQUENCE.	
RC STRAIN=ATCC 43456;	
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;	
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,	
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;	
RT "The genetic basis for the variation in the lipo-oligosaccharide of	
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated	
RT ganglioside mimics in the core oligosaccharide.";	
RL J. Biol. Chem. 277:327-337(2002).	
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms	
CC Distributed under the Creative Commons Attribution-NoDerivs License	
CC -----	
CC ENBL; AF401528; AAL05990.1; -; Genomic_DNA.	
DR SNR; Q93D05; 1-258.	
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.	
DR InterPro; IPR009251; CST-1.	
DR Pfam; PF06002; CST-1; 1.	
KW Glycosyltransferase; Transferase.	
SQ SEQUENCE 291 AA; 34641 MW; F294A04ACBA82882 CRC64;	
Query Match	100.0%; Score 1564; DB 2; Length 291;
Best Local Similarity	100.0%; Pred. No. 2.7e-104;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCNQFYFDKYLKKCKTVFTYTNPFPEQY 60	
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCNQFYFDKYLKKCKTVFTYTNPFPEQY 60	
QY 61 TLKHLTONQBYETELMCSNYNOAHLENENFVKTFYDFDAHLGYDFPKLKEFNAYFK 120	
Db 61 TLKHLTONQBYETELMCSNYNOAHLENENFVKTFYDFDAHLGYDFPKLKEFNAYFK 120	
QY 121 FHEIYNQRTTGSVYWCVAIAALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180	
Db 121 FHEIYNQRTTGSVYWCVAIAALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180	
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240	
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240	



```
QY 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291
|||||
Db 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

RESULT 2
Q9L905_CAMJE
ID Q9L905 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q9L905;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 21-FEB-2006, entry version 18.
DE Alpha-2,3-sialyltransferase.
GN Name=cst-II; Synonyms=cst, cstII;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43446;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brisson J.-R., Karwaski M.-F., Michniewicz J.,
RA Cunningham A.-M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT MHz (1)H and (13)C NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43446;
RX PubMed=14742567; DOI=10.1128/IAI.70.2.787-793.2002;
RA Gilbert M.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=81-176;
RX MEDLINE=21655167; PubMed=11796612; DOI=10.1128/IAI.70.2.787-793.2002;
RA Guerry P., Szymanski C.M., Prendergast M.M., Hickey T.E., Ewing C.P.,
RA Pattarini D.L., Moran A.P.;
RT "Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide
RT affects ganglioside mimicry and invasiveness in vitro.";
RL Infect. Immun. 70:787-793(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GB11;
RX PubMed=14742567; DOI=10.1128/IAI.72.2.1162-1165.2004;
RA Gilbert M., Godschalk P.C., Karwaski M.F., Ang C.W., Van Belkum A.,
RA Li J., Wakarchuk W.W., Endtz H.P.;
RT "Evidence for acquisition of the lipooligosaccharide biosynthesis
RT locus in Campylobacter jejuni GB11, a strain isolated from a patient
RT with Guillain-Barre syndrome, by horizontal exchange.";
RL Infect. Immun. 72:1162-1165(2004).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CF90-26;
RX Koga M., Gilbert M., Takahashi M., Li J., Koike S., Hirata K.,
RA Yuki N.;
RT "Comprehensive Analysis of Bacterial Risk Factors for the Development
RT of Guillain-Barre Syndrome after Campylobacter jejuni Enteritis.";
RL J. Infect. Dis. 193:547-555(2006).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF167344; AAF34137.1; -; Genomic DNA.
DR EMBL; AF305571; AAL09368.1; -; Genomic DNA.
DR EMBL; AY422197; AAB82875.1; -; Genomic DNA.
DR EMBL; AY661458; AAV80723.1; -; Genomic DNA.
DR SMR; Q9L905; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.

DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34611 MW; 1BE4A04ACD728228 CRC64;

Query Match 99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 6.1e-104;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRLPNDPFRNCNQFYFEDKYLGGKCKTVFTPTNFFRQYY 60
|||||
Db 1 MKKVIAGNGPSLKEIDYSRLPNDPFRNCNQFYFEDKYLGGKCKAVFTPTNFFRQYY 60
|||||
QY 61 TLKHLIQOEYETELIMCSYNAHLNENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
|||||
Db 61 TLKHLIQOEYETELIMCSYNAHLNENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
|||||
QY 121 FHEIYFNORITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
|||||
Db 121 FHEIYFNORITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
|||||
QY 181 DRSHVIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
|||||
Db 181 DRSHVIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
|||||
QY 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291
|||||
Db 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291
|||||

RESULT 3
Q9F0M9_CAMJE
ID Q9F0M9 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q9F0M9;
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Alpha-2,3-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43432;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF215659; AAG43979.1; -; Genomic DNA.
DR SMR; Q9F0M9; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34554 MW; E41B594ACD7280F8 CRC64;

Query Match 99.3%; Score 1553; DB 2; Length 291;
Best Local Similarity 99.3%; Pred. No. 1.6e-103;
Matches 289; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRLPNDPFRNCNQFYFEDKYLGGKCKTVFTPTNFFRQYY 60
|||||
Db 1 MKKVIAGNGPSLKEIDYSRLPNDPFRNCNQFYFEDKYLGGKCKAVFTPTNFFRQYY 60
|||||
QY 61 TLKHLIQOEYETELIMCSYNAHLNENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
|||||
Db 61 TLKHLIQOEYETELIMCSYNAHLNENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
|||||
QY 121 FHEIYFNORITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
|||||
DR InterPro; IPR009251; CST-I.
```

```
Db 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Qy 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Qy 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 4
Q93M00_CAMJE
ID Q93M00_CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93M00;
DT 01-DEC-2001, integrated into UniProtKB/TREMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Alpha-2,3-/alpha-2,8-sialyltransferase.
GN Name=cstii;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43438;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF400048; AAK91725.1; -; Genomic_DNA.
DR SMR; Q93M00; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34573 MW; BDIEDEBF6F521E6 CRC64;

Query Match 97.3%; Score 1521; DB 2; Length 291;
Best Local Similarity 97.3%; Pred. No. 3.3e-101;
Matches 283; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MKKVIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIISNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDFPDALHGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDFPDALHGYDFKQKKEFNAYFK 120
Qy 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Qy 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Qy 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 5
Q9LAK3_CAMJE
ID Q9LAK3_CAMJE PRELIMINARY; PRT; 291 AA.
AC Q9LAK3;
DT 01-OCT-2000, integrated into UniProtKB/TREMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Alpha-2,3-/8-sialyltransferase (Alpha-2,3-/2,8-sialyltransferase).
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OH4384;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brissou J.-R., Karwaski M.-F., Michniewicz J.,
RA Cunningham A.-M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT MHz (1H) and (13C) NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 700297;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CF 93-6;
RA Koga M., Gilbert M., Li J., Koike S., Takahashi M., Furukawa K.,
RA Hirata K., Yuki N.;
RT "Antecedent infections in Fisher Syndrome: a common pathogenesis of
RT molecular mimicry";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF130984; AAF31771.1; -; Genomic_DNA.
DR EMBL; AF216647; AAL36462.1; -; Genomic_DNA.
DR EMBL; AY644679; AAV52039.1; -; Genomic_DNA.
DR PDB; 1RO7; X-ray; A/B/C/D=1-259.
DR PDB; 1RO8; X-ray; A/B=1-259.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34545 MW; AB0741D5FCEDAE6A CRC64;

Query Match 97.3%; Score 1521; DB 2; Length 291;
Best Local Similarity 96.9%; Pred. No. 3.3e-101;
Matches 282; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MKKVIAGNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIISNGPSLKEIDYSLPNDVDFRCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Qy 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDFPDALHGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDFPDALHGYDFKQKKEFNAYFK 120
Qy 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Qy 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Qy 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
```

```
RESULT 6
O50F20 CAMJE PRELIMINARY; PRT; 284 AA.
AC O50F20;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Cj81-011 (Fragment).
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=81-176;
RX PubMed=15872262; DOI=10.1128/JCM.43.5.2330-2338.2005;
RA Poly F., Threagill D., Stintzi A.;
RT "Genomic Diversity in Campylobacter jejuni: Identification of C.
RT jejuni 81-176-Specific Genes.";
RL J. Clin. Microbiol. 43:2330-2338(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY681242; AAW56104.1; -; Genomic_DNA.
DR SMR; O50F20; 1-258.
DR NON TER 284
SQ SEQUENCE 284 AA; 33722 MW; 785BEA268C691E70 CRC64;

Query Match 97.0%; Score 1517; DB 2; Length 284;
Best Local Similarity 99.6%; Pred. No. 6.2e-101;
Matches 283; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNSGLKEIDYSRPNDFVPRCNOQFYEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNSGLKEIDYSRPNDFVPRCNOQFYEDKYLGKKCKAVFTPNFFEQYY 60

Qy 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120

Qy 121 PHEIFYNQRTISGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 PHEIFYNQRTISGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDI 284
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDI 284

RESULT 7
Q938X6 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q938X6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Alpha-2,3/-alpha-2,8-sialyltransferase.
GN Name=cstII;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43460;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;

Qy 1 MKKVIIAGNSGLKEIDYSRPNDFVPRCNOQFYEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNSGLKEIDYSRPNDFVPRCNOQFYEDKYLGKKCKAVFTPNFFEQYY 60

Qy 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120

Qy 121 PHEIFYNQRTISGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 PHEIFYNQRTISGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDI 284
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDI 284

RESULT 8
Q93CZ5 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93CZ5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Bifunctional alpha-2,3/-2,8-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF401529; AAL06004.1; -; Genomic_DNA.
DR SMR; Q93CZ5; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
```

```
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY044868; AAK96001.1; -; Genomic_DNA.
DR SMR; Q938X6; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-1.
DR Pfam; PF06002; CST-1; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34577 MW; 7BAE5F6021A56F08 CRC64;

Query Match 96.4%; Score 1507; DB 2; Length 291;
Best Local Similarity 95.2%; Pred. No. 3.3e-100;
Matches 277; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNSGLKEIDYSRPNDFVPRCNOQFYEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNSGLKEIDYSRPNDFVPRCNOQFYEDKYLGKKCKAVFPNPSLFFEQYY 60

Qy 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120

Qy 121 PHEIFYNQRTISGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 PHEIFYNQRTISGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db 181 DSNHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFTKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
Q93CZ5 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93CZ5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Bifunctional alpha-2,3/-2,8-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF401529; AAL06004.1; -; Genomic_DNA.
DR SMR; Q93CZ5; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
```

```

DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34505 MW; 50162DE8D81558F CRC64;

Query Match
Best Local Similarity 95.4%; Score 1492; DB 2; Length 291;
Matches 274; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQYFEDKYLGKKCKTVFTNPFPEQYY 60
DB 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQYFEDKYLGKKCKAVFYNPSLFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 9
Q5DT12 CAMJE
ID Q5DT12 CAMJE PRELIMINARY; PRT; 292 AA.
AC Q5DT12;
DT 29-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 29-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Putative alpha-2,3/-2,8 sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MF6;
RA Gilbert M., Karwaski M.-F., Godschalk P.C.R., Brochu D., Endtz H.P.,
RA Cunningham A.;
RT "Sequencing of the lipooligosaccharide biosynthesis locus of
RT Campylobacter jejuni MF6.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AY422196; AAR82857.1; -; Genomic DNA.
DR SMR; Q5DT12; 1-259.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 292 AA; 34658 MW; 1AADF7B6FF7B0A2F CRC64;

Query Match
Best Local Similarity 92.8%; Score 1451.5; DB 2; Length 292;
Matches 268; Conservative 13; Mismatches 10; Indels 1; Gaps 1;

QY 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQYFEDKYLGKKCKTVFTNPFPEQYY 60
DB 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQYFEDKYLGKKCKAVFYNPSLFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 10
Q2TKA8 CAMJE
ID Q2TKA8 CAMJE PRELIMINARY; PRT; 263 AA.
AC Q2TKA8;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Putative alpha-2,3/8-sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GB5;
RA Gilbert M., Brochu D., Karwaski M.-F.;
RT "Sequencing of the Campylobacter jejuni GB5 LOS locus.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AY854153; AAX45339.1; -; Genomic DNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 263 AA; 31295 MW; 21E0289B1B9823E0 CRC64;

Query Match
Best Local Similarity 83.5%; Pred. No. 2.3e-77;
Matches 223; Conservative 14; Mismatches 17; Indels 13; Gaps 2;

QY 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQYFEDKYLGKKCKTVFTNPFPEQYY 60
DB 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQYFEDKYLGKKCKAVFYNPSLFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240

QY 241 KDI-----LIPSSAYGKFSKNINFKKI 263
DB 241 LKIYSVLLVRLME-----NFQKI 258

RESULT 11
Q7BP25 CAMJE
ID Q7BP25 CAMJE PRELIMINARY; PRT; 294 AA.
AC Q7BP25; Q9PNF4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.

```

```
DE Hypothetical protein Cj1140 (2,3-sialyl transferase) (Alpha-2,3-
DE sialyltransferase).
DE Name=cstIII; Synonyms=cst-III; OrderedLocusNames=Cj1140;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_taxid=197;
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MSC57360;
RA Guerry P., Ewing C.P., Moran A.P., Trust T.J.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RW
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MSC 57360;
RX MEDLINE=20536403; PubMed=11083778;
RX DOI=10.1128/JAI.68.12.6656-6662.2000;
RA Guerry P., Ewing C.P., Hickey T.E., Prendergast M.M., Moran A.P.;
RT "Sialylation of lipooligosaccharide cores affects immunogenicity and
RT serum resistance of Campylobacter jejuni.";
RL Infect. Immun. 68:6656-6662(2000).
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43429, and ATCC 43430;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCCT 11168;
RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Kately J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajadream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AF195055; AAG29922.1; -; Genomic DNA.
DR EMBL; AF257460; AAG37020.1; -; Genomic DNA.
DR EMBL; AF400047; AAK85419.1; -; Genomic DNA.
DR EMBL; AL139077; CAB73395.1; -; Genomic DNA.
DR EMBL; AY044156; AAK73183.1; -; Genomic DNA.
DR PIR; G81318; G81318.
DR Biocyc; CUEJ192222:Cj1140-MONOMER; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
DR InterPro; IPR009251; CST-1.
DR Pfam; PF06002; CST-1; 1.
KW Complete proteome; Glycosyltransferase; Hypothetical protein;
KW Transferrase.
SQ SEQUENCE 294 AA; 35127 MW; 1729C9C35797D78F CRC64;

Query Match 51.4%; Score 804; DB 2; Length 294;
Best Local Similarity 53.9%; Pred. No. 1,1e-49;
Matches 158; Conservative 45; Mismatches 77; Indels 18; Gaps 6;

QY 5 IIAAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKTFTYTPNPFPEQYTLKH 64
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
8 LVCGNPSLKNIDYKRLKQDFVRCNQFYFEDRYFVGKDYKVFNFVFPFPEQYTSKK 67
QY 65 LIQOEYETELIMCSNYNQAHLNENFVKTYDYPDDAHLGYDFPKQKFNAYFKFHEI 124
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
68 LIQNEEYNIENICVSTINLEYIDQFVDNPFELVSDAFLGHEIITKLDFFAYIKYNEI 127
```

```
QY 125 YFNQRITSGVYMCVAIAIAGLYKEIYLSGIDFYQNGSS-YAFDTKQENLLKLAPDFKDRS 183
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
128 YNRQRITSGVYMCATVALGYSIYSGIDFYQTNLLYAFDNKKLLNKCTGFKQKF 187
QY 184 HYGHSKNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLNSNFIQEK-NNTYKD 242
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
188 KFINHSMACDLQALDYMKRYDVIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
QY 243 ILIPSEA----YGFPSKINPKIKIKENVYKLIKDLRLPSDIKHYFKGK 291
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
244 ILIPDKYAQERYYG-----KSLKENLHYKLIKDLRLPSDIKHYLKEK 288

RESULT 12
Q4QM36_HAE18 PRELIMINARY; PRT; 326 AA.
AC Q4QM36;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE CMP-neu5AC--lipooligosaccharide alpha 2-3 sialyltransferase.
GN Name=lic3A2; OrderedLocusNames=NTHI1034;
OS Haemophilus influenzae (strain 86-028NP).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=281310;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA]
RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;
RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,
RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,
RA Munson R.S. Jr.;
RT "Genomic sequence of an otitis media isolate of nontypeable
RT Haemophilus influenzae: comparative study with H. influenzae serotype
RT d, strain Kw20.";
RL J. Bacteriol. 187:4627-4636(2005).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; CP000057; AX87911.1; -; Genomic DNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
KW Complete proteome; Glycosyltransferase; Transferrase.
SQ SEQUENCE 326 AA; 38272 MW; 78FF283EB10C6F0E CRC64;

Query Match 47.6%; Score 744.5; DB 2; Length 326;
Best Local Similarity 50.7%; Pred. No. 2.3e-45;
Matches 149; Conservative 50; Mismatches 86; Indels 9; Gaps 6;

QY 2 KKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKTFTYTPNPFPEQYTT 61
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
32 KSVIIAGNPSLKSIDYSLPKDYDVRFCNQFYFEDHYFLGKIKKVFVNCVIFEQYTT 91
QY 62 LKHLIQOEYE-TELIMCSNYNQAHLNENFVKTYDYPDDAHLGYDFPKQKFNAYFK 120
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
92 FMOLIKNNEYADIIISLFSNLGDSLKK-IQRLKLLPQIDLGHSYLLKLRADFADHLQ 150
QY 121 FHEIYFNQIRITSGVYMCVAIAIAGLYKEIYLSGIDFYQ-NGSSYAFDTKQENLLKLPDFK 179
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
151 YHELYENKRITSGVYMCVAVATAMGYKDYLTGDFYQKGNPYAFHQTNIIKLPSFS 210
QY 180 NDRSHVIGHSKNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLNS--NFIQEK 237
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
211 QKQSDSIHSMEDVNLALYFLQKHGVNIYICISPSPLCMYFPLSP-LNNPITFLEKK 269
QY 238 NYTKDILIPSSBAYGKFSKNINFKIKIKENVYKLIKDLRLPSDIKHYFKGK 291
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
270 NYTDILIPPKFYKIG---IYSPRIYQNLIFRLIWDILRLPNDIKHAKAK 320

RESULT 13
Q9RGF1_CAMJE PRELIMINARY; PRT; 430 AA.
ID Q9RGF1_CAMJE
```

AC O9RGF1;  
 DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.  
 DT 01-MAY-2000, sequence version 1.  
 DT 07-FEB-2006, entry version 16.  
 DE Alpha-2,3-sialyltransferase.  
 GN Name=cst-I;  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=OH4384;  
 RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;  
 RA Gilbert M., Brisson J.-R., Karwaski M.-F., Michniewicz J.,  
 RA Cunningham A.-M., Wu Y., Young N.M., Wakarchuk W.W.;  
 RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.  
 RT Identification of the glycosyltransferase genes, enzymatic synthesis  
 RT of model compounds, and characterization of nanomole amounts by 600-  
 RT MHz (1)H and (13)C NMR analysis.";  
 RN J. Biol. Chem. 275:3896-3906(2000).  
 RL [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=OH4384;  
 RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;  
 RA Gilbert M., Wakarchuk W.W.;  
 RA "Lipopolysaccharide alpha-2,3 sialyltransferase of campylobacter.";  
 RT Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DE EMBL; AF130466; AAF13495.1; -; Genomic DNA.  
 DE GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.  
 DE InterPro; IPR009251; CST-I.  
 DE Pfam; PF06002; CST-I; 1.  
 DE Glycosyltransferase; Transferase.  
 DE SEQUENCE 430 AA; 50396 MW; 1D03B6797169425C CRC64;  
 SQ  
 Query Match 47.3%; Score 739; DB 2; Length 430;  
 Best Local Similarity 50.2%; Pred. No. 7.9e-45;  
 Matches 146; Conservative 45; Mismatches 82; Indels 18; Gaps 5;  
 1 MKKVIAGNGPSLKEIDYSRLPNDPFRVRCNQFYEDKYYLGGKCKTVYTPNFFPEQYY 60  
 16 MQNIIAGNGPSLKNINYKRLPREYDVRFCNQFYEDKYYLGGKCKTVYTPNFFPEQYY 75  
 61 TLKHLIQEQYETELIMCSNYQAHLNENFVKTFDYDPPDAHLGYDFFKQLKEFNAYFK 120  
 76 TAKQILKNEYEIKNIFCSFNLFFIESNDFLHQFYNFPPDAKLGVEVLENLKEFYAYIK 135  
 121 FHEIFYNQRITSGVYMCVAIAALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAP--D 177  
 136 YNEIFYNKRITSGVYMCVAIAALGYKTYILCGIDFYEGDVYVPEAMSTNITKTFPGIKD 195  
 178 FKNDRSHVIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK- 236  
 196 FKPSNC---HSKEYDIEALKLSYKVNIVYALCDDSIILNHFPLSINNNTLENKH 251  
 237 NNYTKDILIPSEAYGKFSKN-----INFKKI-KIKENVYKLIKDL 277  
 252 NNSINDILLTDNTPGVSYFNQKADNKMILNFMNLFILHSDNLIKFLNKEI 302  
 RESULT 14  
 Q32VR8 CAMJE PRELIMINARY; PRT; 430 AA.  
 ID Q32VR8\_CAMJE  
 AC Q32VR8;  
 DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.  
 DT 07-FEB-2006, entry version 1.  
 DE Alpha-2,3-sialyltransferase.  
 GN Name=cstI;  
 OS Campylobacter jejuni subsp. jejuni.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=32022;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 43446;  
 RA Gilbert M., Brochu D., Karwaski M.-F.;  
 RT "Sequencing of the Cst-I locus of Campylobacter jejuni ATCC 43446.";  
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DE EMBL; AY791515; AAX33819.1; -; Genomic DNA.  
 DE GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.  
 DE Glycosyltransferase; Transferase.  
 DE SEQUENCE 430 AA; 50383 MW; 876F9E0583694240 CRC64;  
 SQ  
 Query Match 47.1%; Score 736.5; DB 2; Length 430;  
 Best Local Similarity 51.9%; Pred. No. 1.2e-44;  
 Matches 147; Conservative 40; Mismatches 87; Indels 9; Gaps 4;  
 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRVRCNQFYEDKYYLGGKCKTVYTPNFFPEQYY 60  
 16 MQNIIAGNGPSLKNINYKRLPREYDVRFCNQFYEDKYYLGGKCKTVYTPNFFPEQYY 75  
 61 TLKHLIQEQYETELIMCSNYQAHLNENFVKTFDYDPPDAHLGYDFFKQLKEFNAYFK 120  
 76 TAKQILKNEYEIKNIFCSFNLFFIESNDFLHQFYNFPPDAKLGVEVLENLKEFYAYIK 135  
 121 FHEIFYNQRITSGVYMCVAIAALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAP--D 177  
 136 YNEIFYNKRITSGVYMCVAIAALGYKTYILCGIDFYEGDVYVPEAMSTNITKTFPGIKD 195  
 178 FKNDRSHVIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEK- 236  
 196 FKPSNC---HSKEYDIEALKLSYKVNIVYALCDDSIILNHFPLSINNNTLENKH 251  
 237 NNYTKDILIPSEAYGKFSKNINFKKIKENVYKLIKDL 278  
 252 NNSINDILLTDNTPGVSYFNQKADNKMILNFMNLFILHSDNLIKFLNKEI 294  
 RESULT 15  
 Q4QN18 HAEI8 PRELIMINARY; PRT; 320 AA.  
 ID Q4QN18\_HAEI8  
 AC Q4QN18;  
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.  
 DT 19-JUL-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 5.  
 DE CMP-Neu5Ac--lipooligosaccharide alpha 2-3 sialyltransferase.  
 GN Name=lic3A; Ordered locus names=NH10472;  
 OS Haemophilus influenzae (strain 86-028NP).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=281310;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;  
 RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,  
 RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,  
 RA Munson R.S. Jr.;  
 RT "Genomic sequence of an otitis media isolate of nontypeable  
 RT Haemophilus influenzae: comparative study with H. influenzae serotype  
 RT d, strain KW20.";  
 RL J. Bacteriol. 187:4627-4636(2005).  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC -----  
 DE EMBL; CP000057; AAX87409.1; -; Genomic DNA.  
 DE GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.  
 DE Complete proteome; Glycosyltransferase; Transferase.

Search completed: July 19, 2006, 13:46:48  
Job time : 139.396 secs

GenCore version 5.1.9  
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:47:11 ; Search time 27.928 Seconds  
(without alignments)  
912.038 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIIAGNGPSLKEIDYR.....KLIKOLLRLPSDIKHVFKGK 291

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA:\*

1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5 COMB.pep.\*

2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6 COMB.pep.\*

3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7 COMB.pep.\*

4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/H COMB.pep.\*

5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/pCTUS COMB.pep.\*

6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE COMB.pep.\*

7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	100.0	291	2	US-09-816-028A-14
2	1564	100.0	291	2	US-10-303-162-14
3	1564	100.0	291	2	US-10-303-134-14
4	1564	100.0	291	2	US-10-303-118-14
5	1564	100.0	291	2	US-10-303-128-14
6	1564	100.0	291	3	US-10-735-419-14
7	1559	99.7	291	2	US-09-495-406-9
8	1559	99.7	291	2	US-09-816-028A-9
9	1559	99.7	291	2	US-10-303-162-9
10	1559	99.7	291	2	US-10-303-134-9
11	1559	99.7	291	2	US-10-303-118-9
12	1559	99.7	291	2	US-10-303-128-9
13	1559	99.7	291	3	US-10-735-419-9
14	1553	99.3	291	2	US-09-816-028A-12
15	1553	99.3	291	2	US-10-303-162-12
16	1553	99.3	291	2	US-10-303-134-12
17	1553	99.3	291	2	US-10-303-118-12
18	1553	99.3	291	2	US-10-303-128-12
19	1553	99.3	291	3	US-10-735-419-12
20	1524	97.4	291	2	US-09-495-406-5
21	1524	97.4	291	2	US-09-816-028A-5
22	1524	97.4	291	2	US-10-303-162-5
23	1524	97.4	291	2	US-10-303-134-5
24	1524	97.4	291	2	US-10-303-118-5
25	1524	97.4	291	2	US-10-303-128-5
26	1524	97.4	291	3	US-10-735-419-5

27	1521	97.3	291	2	US-09-495-406-3	Sequence 3, Appli
28	1521	97.3	291	2	US-09-816-028A-3	Sequence 3, Appli
29	1521	97.3	291	2	US-10-303-162-3	Sequence 3, Appli
30	1521	97.3	291	2	US-10-303-134-3	Sequence 3, Appli
31	1521	97.3	291	2	US-10-303-118-3	Sequence 3, Appli
32	1521	97.3	291	2	US-10-303-128-3	Sequence 3, Appli
33	1521	97.3	291	2	US-10-735-419-3	Sequence 3, Appli
34	1510	96.5	291	2	US-09-495-406-7	Sequence 7, Appli
35	1510	96.5	291	2	US-09-816-028A-7	Sequence 7, Appli
36	1510	96.5	291	2	US-10-303-162-7	Sequence 7, Appli
37	1510	96.5	291	2	US-10-303-134-7	Sequence 7, Appli
38	1510	96.5	291	2	US-10-303-118-7	Sequence 7, Appli
39	1510	96.5	291	2	US-10-303-128-7	Sequence 7, Appli
40	1510	96.5	291	3	US-10-735-419-7	Sequence 7, Appli
41	804	51.4	294	2	US-09-495-406-10	Sequence 10, Appl
42	804	51.4	294	2	US-09-816-028A-10	Sequence 10, Appl
43	804	51.4	294	2	US-10-303-162-10	Sequence 10, Appl
44	804	51.4	294	2	US-10-303-134-10	Sequence 10, Appl
45	804	51.4	294	2	US-10-303-118-10	Sequence 10, Appl

## ALIGNMENTS

### RESULT 1

US-09-816-028A-14

; Sequence 14, Application US/09816028A

; Patent No. 6699705

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-000111US

; CURRENT APPLICATION NUMBER: US/09/816,028A

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/118,213

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406

; PRIOR FILING DATE: 2000-01-31

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 291

; TYPE: PRT

; ORGANISM: Campylobacter jejuni

; FEATURE:

; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II

; OTHER INFORMATION: (CstII) from C. jejuni O:36

US-09-816-028A-14

Query Match 100.0%; Score 1564; DB 2; Length 291;

Best Local Similarity 100.0%; Pred. No. 7e-154;

Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYRPNDFVRCNQVFEDKYLGKKCTVFTYTPNFFEEQY 60

Db 1 MKKVIIAGNGPSLKEIDYRPNDFVRCNQVFEDKYLGKKCTVFTYTPNFFEEQY 60

QY 61 TLKHLIQNOEYETELIMCSNQAHLNENFVKTFDYFPDAHLGYDFFKOLKEFNAYFK 120

Db 61 TLKHLIQNOEYETELIMCSNQAHLNENFVKTFDYFPDAHLGYDFFKOLKEFNAYFK 120

QY 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTTKOENLLKLAPDFKN 180

Db 121 FHEIFYNQRTSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTTKOENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKYCLCPNSLLANFIELAPLNSNFIIQEKNNYT 240

Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKYCLCPNSLLANFIELAPLNSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINPKIKENVYVKIKIDLLRLPSDIKHVFKGK 291



Db 241 KDILPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFGK 291

RESULT 2

US-10-303-162-14  
 ; Sequence 14, Application US/10303162  
 ; Patent No. 6723545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/303,162  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-303-162-14

Query Match 100.0%; Score 1564; DB 2; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 7e-154;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNQYFEDKYLGGKCKTVFTPNRFFQYY 60  
 Db 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNQYFEDKYLGGKCKTVFTPNRFFQYY 60  
 QY 61 TLKHLIQNOYETELIMCSYNOAHLENENFVKTFYDYPDAHLGYDFFKQKFNAYFK 120  
 Db 61 TLKHLIQNOYETELIMCSYNOAHLENENFVKTFYDYPDAHLGYDFFKQKFNAYFK 120  
 QY 121 FHEIFYNQRITSGVYMCVAIAIYALGKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 Db 121 FHEIFYNQRITSGVYMCVAIAIYALGKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHVIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHVIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFGK 291  
 Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFGK 291

RESULT 3

US-10-303-134-14  
 ; Sequence 14, Application US/10303134  
 ; Patent No. 6825019  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/303,134  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028

; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-303-134-14

Query Match 100.0%; Score 1564; DB 2; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 7e-154;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNQYFEDKYLGGKCKTVFTPNRFFQYY 60  
 Db 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNQYFEDKYLGGKCKTVFTPNRFFQYY 60  
 QY 61 TLKHLIQNOYETELIMCSYNOAHLENENFVKTFYDYPDAHLGYDFFKQKFNAYFK 120  
 Db 61 TLKHLIQNOYETELIMCSYNOAHLENENFVKTFYDYPDAHLGYDFFKQKFNAYFK 120  
 QY 121 FHEIFYNQRITSGVYMCVAIAIYALGKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 Db 121 FHEIFYNQRITSGVYMCVAIAIYALGKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHVIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 Db 181 DRSHVIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFGK 291  
 Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHFGK 291

RESULT 4

US-10-303-118-14  
 ; Sequence 14, Application US/10303118  
 ; Patent No. 6905867  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/303,118  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-303-118-14

Query Match 100.0%; Score 1564; DB 2; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 7e-154;

Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNPSLKEIDYSLPNDVFRNCQYFEDKYYLGKCKTVFTPNFFEQYY 60  
DB 1 MKKVIIAGNPSLKEIDYSLPNDVFRNCQYFEDKYYLGKCKTVFTPNFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120  
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIYFNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 FHEIYFNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKIDLLRLPSDIKHYPKKG 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKIDLLRLPSDIKHYPKKG 291

## RESULT 5

US-10-303-128-14  
; Sequence 14, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-128-14

Query Match 100.0%; Score 1564; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 7e-154;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNPSLKEIDYSLPNDVFRNCQYFEDKYYLGKCKTVFTPNFFEQYY 60  
DB 1 MKKVIIAGNPSLKEIDYSLPNDVFRNCQYFEDKYYLGKCKTVFTPNFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120  
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIYFNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 FHEIYFNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKIDLLRLPSDIKHYPKKG 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKIDLLRLPSDIKHYPKKG 291

RESULT 6  
US-10-735-419-14  
; Sequence 14, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-735-419-14

Query Match 100.0%; Score 1564; DB 3; Length 291;  
Best Local Similarity 100.0%; Pred. No. 7e-154;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNPSLKEIDYSLPNDVFRNCQYFEDKYYLGKCKTVFTPNFFEQYY 60  
DB 1 MKKVIIAGNPSLKEIDYSLPNDVFRNCQYFEDKYYLGKCKTVFTPNFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120  
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIYFNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 FHEIYFNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKIDLLRLPSDIKHYPKKG 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKIDLLRLPSDIKHYPKKG 291

## RESULT 7

US-09-495-406-9  
; Sequence 9, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; CURRENT FILING DATE: 2000-01-31

```

; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
US-09-495-406-9

```

```

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.3e-153;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRPNDFVRCNOFYEDKYYLGGKCKTFTVTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVRCNOFYEDKYYLGGKCKTFTVTPNFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSYNOAHLENNFVKTFTYDYPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSYNOAHLENNFVKTFTYDYPDAHLGYDFFKQKKEFNAYFK 120

QY 121 FHEIYFNQIRITSGVYMCVAIAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYFNQIRITSGVYMCVAIAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYFKGK 291

```

```

RESULT 8
US-09-816-028A-9
; Sequence 9, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 015633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 015633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-816-028A-9

```

```

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.3e-153;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRPNDFVRCNOFYEDKYYLGGKCKTFTVTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVRCNOFYEDKYYLGGKCKTFTVTPNFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSYNOAHLENNFVKTFTYDYPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSYNOAHLENNFVKTFTYDYPDAHLGYDFFKQKKEFNAYFK 120

QY 121 FHEIYFNQIRITSGVYMCVAIAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYFNQIRITSGVYMCVAIAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYFKGK 291

```

```

RESULT 10
US-10-303-134-9
; Sequence 9, Application US/10303134

```

```
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-134-9

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.3e-153;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRPNDFVFCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVFCNQFYFEDKYLGKKCKAVFTPNFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFPKOLFKNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFPKOLFKNAYFK 120

QY 121 FHEIFYNQRITSGVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180
Db 121 FHEIFYNQRITSGVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYFKGK 291

RESULT 11
US-10-303-118-9
; Sequence 9, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-118-9

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.3e-153;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRPNDFVFCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVFCNQFYFEDKYLGKKCKAVFTPNFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFPKOLFKNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFPKOLFKNAYFK 120

QY 121 FHEIFYNQRITSGVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180
Db 121 FHEIFYNQRITSGVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYFKGK 291

RESULT 12
US-10-303-128-9
; Sequence 9, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-128-9

Query Match          99.7%; Score 1559; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.3e-153;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRPNDFVFCNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRPNDFVFCNQFYFEDKYLGKKCKAVFTPNFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFPKOLFKNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFPKOLFKNAYFK 120

QY 121 FHEIFYNQRITSGVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180
Db 121 FHEIFYNQRITSGVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYFKGK 291
```

```

Db 61 TLKHLIQEYETELIMCNSYNQAHLNENFVKTFYDFPDHLYGDFKQLKEFNAYFK 120
QY 121 FHEIYFNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYFNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

```

```

RESULT 13
US-10-735-419-9
; Sequence 9, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-735-419-9

```

```

Query Match 99.7%; Score 1559; DB 3; Length 291;
Best Local Similarity 99.7%; Pred. No. 2.3e-153;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRCNQYFEDKYLGKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRCNQYFEDKYLGKCKTVFTPNFFEQYY 60
QY 61 TLKHLIQEYETELIMCNSYNQAHLNENFVKTFYDFPDHLYGDFKQLKEFNAYFK 120
Db 61 TLKHLIQEYETELIMCNSYNQAHLNENFVKTFYDFPDHLYGDFKQLKEFNAYFK 120
QY 121 FHEIYFNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYFNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

RESULT 14
US-09-816-028A-12

```

```

; Sequence 12, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-09-816-028A-12

Query Match 99.3%; Score 1553; DB 2; Length 291;
Best Local Similarity 99.3%; Pred. No. 9.7e-153;
Matches 289; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRCNQYFEDKYLGKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRCNQYFEDKYLGKCKTVFTPNFFEQYY 60
QY 61 TLKHLIQEYETELIMCNSYNQAHLNENFVKTFYDFPDHLYGDFKQLKEFNAYFK 120
Db 61 TLKHLIQEYETELIMCNSYNQAHLNENFVKTFYDFPDHLYGDFKQLKEFNAYFK 120
QY 121 FHEIYFNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYFNORITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

```

```

RESULT 15
US-10-303-162-12
; Sequence 12, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12

```

```
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-162-12

Query Match      99.3%; Score 1553; DB 2; Length 291;
Best Local Similarity 99.3%; Pred. NO. 9.7e-153;
Matches 289; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 MKKVIAGNGPSLKEIDYSRLPNDEDFVRCNQFYEDKYILGKKCKTVYTPNFFFEQYY 60
Db      1 MKKVIAGNGPSLKEIDYSRLPNDEDFVRCNQFYEDKYILGKKCKTVYTPNFFFEQYY 60

QY      61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYFPDAHLGYDFFKQKLFNAYFK 120
Db      61 TLKHLIQNOEYETELIMCSNYNOAHLENENFVKTFYDYFPDAHLGYDFFKQKLFNAYFK 120

QY      121 PHEIYFNQRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFFKN 180
Db      121 PHEIYFNQRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFFKN 180

QY      181 DRSHYIGHSKNTDIIKALEPLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIIKALEPLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY      241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLPLSDIKHYFKGK 291
Db      241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLPLSDIKHYFKGK 291
```

Search completed: July 19, 2006, 13:50:03  
Job time : 28.928 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 14:13:03 ; Search time 95.0051 Seconds  
(without alignments)  
1418.823 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIAGNGSLKEIDYSR.....KLIKDLLRLPSDIKHVFKGK 291

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1564	100.0	291	3	US-09-816-028A-14
2	1564	100.0	291	4	US-10-303-161-14
3	1564	100.0	291	4	US-10-303-118-14
4	1564	100.0	291	4	US-10-303-128-14
5	1564	100.0	291	4	US-10-303-134-14
6	1564	100.0	291	4	US-10-303-162-14
7	1564	100.0	291	4	US-10-820-536-14
8	1564	100.0	291	4	US-10-845-408-14
9	1564	100.0	291	4	US-10-845-412-14
10	1564	100.0	291	5	US-10-846-219-14
11	1564	100.0	291	5	US-10-821-604-14
12	1564	100.0	291	5	US-10-847-983-14
13	1564	100.0	291	5	US-10-821-573-14
14	1564	100.0	291	5	US-10-850-807-14
15	1564	100.0	291	5	US-10-850-125-14
16	1564	100.0	291	5	US-10-830-825-14
17	1564	100.0	291	5	US-10-962-334-14
18	1564	100.0	291	5	US-10-830-997-14
19	1564	100.0	291	5	US-10-962-235-14
20	1564	100.0	291	5	US-10-961-882-14
21	1559	99.7	291	3	US-09-816-028A-9
22	1559	99.7	291	4	US-10-303-161-9
23	1559	99.7	291	4	US-10-303-118-9
24	1559	99.7	291	4	US-10-303-128-9
25	1559	99.7	291	4	US-10-303-134-9
26	1559	99.7	291	4	US-10-303-162-9
27	1559	99.7	291	4	US-10-820-536-9

28	1559	99.7	291	4	US-10-845-408-9	Sequence 9, Appli
29	1559	99.7	291	4	US-10-845-412-9	Sequence 9, Appli
30	1559	99.7	291	5	US-10-846-219-9	Sequence 9, Appli
31	1559	99.7	291	5	US-10-821-604-9	Sequence 9, Appli
32	1559	99.7	291	5	US-10-847-983-9	Sequence 9, Appli
33	1559	99.7	291	5	US-10-821-573-9	Sequence 9, Appli
34	1559	99.7	291	5	US-10-850-807-9	Sequence 9, Appli
35	1559	99.7	291	5	US-10-850-125-9	Sequence 9, Appli
36	1559	99.7	291	5	US-10-830-825-9	Sequence 9, Appli
37	1559	99.7	291	5	US-10-962-334-9	Sequence 9, Appli
38	1559	99.7	291	5	US-10-830-997-9	Sequence 9, Appli
39	1559	99.7	291	5	US-10-962-235-9	Sequence 9, Appli
40	1559	99.7	291	5	US-10-961-882-9	Sequence 9, Appli
41	1553	99.3	291	3	US-09-816-028A-12	Sequence 12, Appli
42	1553	99.3	291	4	US-10-303-161-12	Sequence 12, Appli
43	1553	99.3	291	4	US-10-303-118-12	Sequence 12, Appli
44	1553	99.3	291	4	US-10-303-128-12	Sequence 12, Appli
45	1553	99.3	291	4	US-10-303-134-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1

US-09-816-028A-14  
; Sequence 14, Application US/09816028A  
; Patent No. US20020042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-09-816-028A-14

Query Match		100.0%;	Score 1564;	DB 3;	Length 291;
Best Local Similarity		100.0%;	Pred. No. 7.5e-128;	Mismatches 0;	Indels 0;
Matches 291;		Conservative 0;			
Qy	1	MKKVIAGNGPSLKEIDYSRLPNDVFRNCNQVFEDKYLGKCKTKVFTPNFFEOYY	60		
Db	1	MKKVIAGNGPSLKEIDYSRLPNDVFRNCNQVFEDKYLGKCKTKVFTPNFFEOYY	60		
Qy	61	TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK	120		
Db	61	TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK	120		
Qy	121	FHEIFVNRITSGVYNCVAIAIALGYKEIYLSGIDFYONGSSYAPDTQENLLKLAPFKN	180		
Db	121	FHEIFVNRITSGVYNCVAIAIALGYKEIYLSGIDFYONGSSYAPDTQENLLKLAPFKN	180		
Qy	181	DRSHYGHGSKNTDIIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240		
Db	181	DRSHYGHGSKNTDIIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240		
Qy	241	KDILIPSSBAYGKFSKNINFKIKIKENVYKLIKIKOLLRLPSDIKHVFKGK	291		
Db	241	KDILIPSSBAYGKFSKNINFKIKIKENVYKLIKIKOLLRLPSDIKHVFKGK	291		

Db 241 KDILIPSEAYGKFSKNINFKKIKIKENYVKLIKOLLRLPSDIKHVFKGK 291

## RESULT 2

US-10-303-161-14  
; Sequence 14, Application US/10303161  
; Publication No. US20030148459A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,161  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-161-14

Query Match 100.0%; Score 1564; DB 4; Length 291;

Best Local Similarity 100.0%; Pred. No. 7.5e-128; Indels 0; Gaps 0;  
Matches 291; Conservative 0; Mismatches 0;

Qy	1	MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQNFYFEDKYLGKKCKTVFTPNFFFEQYY	60
Db	1	MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQNFYFEDKYLGKKCKTVFTPNFFFEQYY	60
Qy	61	TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK	120
Db	61	TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK	120
Qy	121	FHEIYNQIRITSGVYMCVAVALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN	180
Db	121	FHEIYNQIRITSGVYMCVAVALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN	180
Qy	181	DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240
Db	181	DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240
Qy	241	KDILIPSEAYGKFSKNINFKKIKIKENYVKLIKOLLRLPSDIKHVFKGK	291
Db	241	KDILIPSEAYGKFSKNINFKKIKIKENYVKLIKOLLRLPSDIKHVFKGK	291

## RESULT 3

US-10-303-118-14  
; Sequence 14, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-118-14

Query Match 100.0%; Score 1564; DB 4; Length 291;

Best Local Similarity 100.0%; Pred. No. 7.5e-128; Indels 0; Gaps 0;  
Matches 291; Conservative 0; Mismatches 0;

Qy	1	MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQNFYFEDKYLGKKCKTVFTPNFFFEQYY	60
Db	1	MKKVIIAGNGPSLKEIDYSRLPNDPFRNCQNFYFEDKYLGKKCKTVFTPNFFFEQYY	60
Qy	61	TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK	120
Db	61	TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK	120
Qy	121	FHEIYNQIRITSGVYMCVAVALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN	180
Db	121	FHEIYNQIRITSGVYMCVAVALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDFKN	180
Qy	181	DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240
Db	181	DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240
Qy	241	KDILIPSEAYGKFSKNINFKKIKIKENYVKLIKOLLRLPSDIKHVFKGK	291
Db	241	KDILIPSEAYGKFSKNINFKKIKIKENYVKLIKOLLRLPSDIKHVFKGK	291

## RESULT 4

US-10-303-128-14  
; Sequence 14, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-128-14

Query Match 100.0%; Score 1564; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 7.5e-128;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 MKKVIIAGNPSLKEIDYSLPNDVDFRCNQYFEDKYLYGKCKTVFTYTNFFFEQYY 60  
Db 1 MKKVIIAGNPSLKEIDYSLPNDVDFRCNQYFEDKYLYGKCKTVFTYTNFFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFFDPAHLGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFFDPAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHFKGK 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHFKGK 291

## RESULT 5

US-10-303-134-14  
; Sequence 14, Application US/10303134  
; Publication No. US20030157657A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-134-14

Query Match 100.0%; Score 1564; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 7.5e-128;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNPSLKEIDYSLPNDVDFRCNQYFEDKYLYGKCKTVFTYTNFFFEQYY 60  
Db 1 MKKVIIAGNPSLKEIDYSLPNDVDFRCNQYFEDKYLYGKCKTVFTYTNFFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFFDPAHLGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFFDPAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHFKGK 291

Db 241 KDILIPSSAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHFKGK 291  
RESULT 6  
US-10-303-162-14  
; Sequence 14, Application US/10303162  
; Publication No. US20030157658A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-162-14

Query Match 100.0%; Score 1564; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 7.5e-128;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNPSLKEIDYSLPNDVDFRCNQYFEDKYLYGKCKTVFTYTNFFFEQYY 60  
Db 1 MKKVIIAGNPSLKEIDYSLPNDVDFRCNQYFEDKYLYGKCKTVFTYTNFFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFFDPAHLGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFFDPAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHFKGK 291  
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHFKGK 291

## RESULT 7

US-10-820-536-14  
; Sequence 14, Application US/10820536  
; Publication No. US20040203103A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128

```
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-820-536-14

Query Match      100.0%; Score 1564; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.5e-128;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCKTVFTPNFFFEQYY 60
Db      1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCKTVFTPNFFFEQYY 60

QY      61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
Db      61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY      121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db      121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY      241 KDILIPSEAYGKFSKNINFKKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db      241 KDILIPSEAYGKFSKNINFKKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 9
US-10-845-412-14
; Sequence 14, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-412-14

Query Match      100.0%; Score 1564; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.5e-128;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCKTVFTPNFFFEQYY 60
Db      1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCKTVFTPNFFFEQYY 60

QY      61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
Db      61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY      121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db      121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY      241 KDILIPSEAYGKFSKNINFKKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db      241 KDILIPSEAYGKFSKNINFKKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
US-10-845-408-14
; Sequence 14, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-845-408-14
```

QY 181 DRSHVGHGKNTDIALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQKKNYT 240  
 Db 181 DRSHVGHGKNTDIALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQKKNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKIKIKENVYKLIKIDLLRLPSDIKHFKGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKIKIKENVYKLIKIDLLRLPSDIKHFKGK 291

RESULT 10  
 US-10-846-219-14  
 ; Sequence 14, Application US/10846219  
 ; Publication No. US20040219638A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/846,219  
 ; CURRENT FILING DATE: 2004-05-14  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-846-219-14

Query Match 100.0%; Score 1564; DB 5; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-128;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIIAGNPSLKEIDYSLPNDVFCNQFYFEDKYLGKCKTGYTPTNFFFEQYY 60  
 Db 1 MKKVIIAGNPSLKEIDYSLPNDVFCNQFYFEDKYLGKCKTGYTPTNFFFEQYY 60  
 QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
 QY 121 FHEIFYNQRTSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPKN 180  
 Db 121 FHEIFYNQRTSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPKN 180  
 QY 181 DRSHVGHGKNTDIALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQKKNYT 240  
 Db 181 DRSHVGHGKNTDIALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQKKNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKIKIKENVYKLIKIDLLRLPSDIKHFKGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKIKIKENVYKLIKIDLLRLPSDIKHFKGK 291

RESULT 11  
 US-10-821-604-14  
 ; Sequence 14, Application US/10821604  
 ; Publication No. US20040229263A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/821,604  
 ; CURRENT FILING DATE: 2004-04-08  
 ; PRIOR APPLICATION NUMBER: 10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-821-604-14  
 Query Match 100.0%; Score 1564; DB 5; Length 291;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-128;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIIAGNPSLKEIDYSLPNDVFCNQFYFEDKYLGKCKTGYTPTNFFFEQYY 60  
 Db 1 MKKVIIAGNPSLKEIDYSLPNDVFCNQFYFEDKYLGKCKTGYTPTNFFFEQYY 60  
 QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
 Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
 QY 121 FHEIFYNQRTSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPKN 180  
 Db 121 FHEIFYNQRTSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPKN 180  
 QY 181 DRSHVGHGKNTDIALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQKKNYT 240  
 Db 181 DRSHVGHGKNTDIALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIQKKNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKIKIKENVYKLIKIDLLRLPSDIKHFKGK 291  
 Db 241 KDILIPSSAYGKFSKNINFKIKIKENVYKLIKIDLLRLPSDIKHFKGK 291

RESULT 12  
 US-10-847-983-14  
 ; Sequence 14, Application US/10847983  
 ; Publication No. US20040229272A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/847,983  
 ; CURRENT FILING DATE: 2004-05-17  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni

```
;
;
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-847-983-14

Query Match      100.0%; Score 1564; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.5e-128;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKCKTVFTPNFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYPDPAHLGYDFFKQKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYPDPAHLGYDFFKQKEFNAYFK 120

Qy 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKOLLRLPSDIKHVFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKOLLRLPSDIKHVFKGK 291

RESULT 13
US-10-821-573-14
; Sequence 14, Application US/10821573
; Publication No. US20040229313A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-821-573-14

Query Match      100.0%; Score 1564; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.5e-128;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKCKTVFTPNFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYPDPAHLGYDFFKQKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYPDPAHLGYDFFKQKEFNAYFK 120

Qy 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKOLLRLPSDIKHVFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKOLLRLPSDIKHVFKGK 291

RESULT 14
US-10-850-807-14
; Sequence 14, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-00011US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-850-807-14

Query Match      100.0%; Score 1564; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.5e-128;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKCKTVFTPNFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNQFYFEDKYLGKKCKTVFTPNFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYPDPAHLGYDFFKQKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSYNOAHLENENFVKTFYDYPDPAHLGYDFFKQKEFNAYFK 120

Qy 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQIRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKOLLRLPSDIKHVFKGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKIKENVYVKLIKOLLRLPSDIKHVFKGK 291

RESULT 15
US-10-850-125-14
; Sequence 14, Application US/10850125
; Publication No. US20040259203A1
```

; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-850-125-14

Query Match 100.0%; Score 1564; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.5e-128;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSLRLENDVDFRCNQFYEDKYLGKKCKTFTVTNFFPEQYY 60
Db |||||
QY 1 MKKVIAGNGPSLKEIDYSLRLENDVDFRCNQFYEDKYLGKKCKTFTVTNFFPEQYY 60
Db |||||

QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTYDYDPPDAHLGYDFEKLKEFNAYEK 120
Db |||||

QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTYDYDPPDAHLGYDFEKLKEFNAYEK 120
Db |||||

QY 121 PHEIYNQRIITSGVTMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db |||||

QY 121 PHEIYNQRIITSGVTMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db |||||

QY 181 DRSHYIGHSKNTDIKALBFLEKTYKIKYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db |||||

QY 181 DRSHYIGHSKNTDIKALBFLEKTYKIKYCLCPNSLLANFIELAPNLNSNFIIOEKNNYT 240
Db |||||

QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHYPKGK 291
Db |||||

QY 241 KDILIPSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHYPKGK 291
Db |||||

Search completed: July 19, 2006, 14:20:06
Job time : 97.0051 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 14:13:57 ; Search time 13.2159 Seconds  
(without alignments)  
1269.788 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKKVIAGNGPSLKEIDYR.....KLIKDLRLPSDKHYFGK 291

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 208217 seqs, 57668156 residues

Total number of hits satisfying chosen parameters: 208217

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New.\*  
1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US09\_NEW\_PUB.pap.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US06\_NEW\_PUB.pap.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US07\_NEW\_PUB.pap.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US08\_NEW\_PUB.pap.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/PCT\_NEW\_PUB.pap.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US10\_NEW\_PUB.pap.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US11\_NEW\_PUB.pap.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US60\_NEW\_PUB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102.5	6.6	651	6	US-10-471-571A-252
2	97.5	6.2	496	6	US-10-471-571A-3348
3	96	6.1	701	6	US-10-471-571A-1902
4	95.5	6.1	468	6	US-10-471-571A-5412
5	93.5	6.0	437	7	US-11-293-697-4502
6	92.5	5.9	955	6	US-10-471-571A-708
7	89	5.7	488	7	US-11-056-355B-48594
8	89	5.7	533	7	US-11-056-355B-48593
9	89	5.7	601	7	US-11-056-355B-48592
10	87	5.6	425	6	US-10-953-349-10851
11	87	5.6	429	6	US-10-953-349-10850
12	87	5.6	456	7	US-11-056-355B-79879
13	87	5.6	478	7	US-11-056-355B-79878
14	87	5.6	521	7	US-11-056-355B-79877
15	87	5.6	971	6	US-10-528-563-7
16	86	5.5	396	6	US-10-449-902-50660
17	85	5.4	528	6	US-10-449-902-44124
18	85	5.4	1337	7	US-11-257-500-29
19	84.5	5.4	863	6	US-10-527-411-36
20	84.5	5.4	1127	6	US-10-527-411-54
21	84.5	5.4	1129	6	US-10-527-411-52
22	84	5.4	452	7	US-11-199-489A-124
23	84	5.4	512	7	US-11-056-355B-91714
24	84	5.4	512	7	US-11-056-355B-95470
25	84	5.4	530	7	US-11-056-355B-91713

26	84	5.4	530	7	US-11-056-355B-95469	Sequence 95469, A
27	84	5.4	659	7	US-11-056-355B-91712	Sequence 91712, A
28	84	5.4	659	7	US-11-056-355B-95468	Sequence 95468, A
29	84	5.4	864	7	US-11-251-208-77	Sequence 77, Appl
30	83.5	5.3	506	6	US-10-449-902-52094	Sequence 52094, A
31	83.5	5.3	735	7	US-11-056-355B-81634	Sequence 81634, A
32	83.5	5.3	843	7	US-11-056-355B-81633	Sequence 81633, A
33	83.5	5.3	1143	6	US-10-630-629-1	Sequence 1, Appl1
34	83.5	5.3	1516	7	US-11-257-500-23	Sequence 23, Appl
35	83	5.3	301	6	US-10-471-571A-1484	Sequence 1484, Ap
36	82.5	5.3	188	6	US-10-471-571A-722	Sequence 722, App
37	82	5.2	348	6	US-10-953-349-10852	Sequence 10852, A
38	82	5.2	394	7	US-11-056-355B-58598	Sequence 58598, A
39	82	5.2	398	7	US-11-056-355B-58597	Sequence 58597, A
40	82	5.2	965	7	US-11-178-560-3	Sequence 3, Appl1
41	81.5	5.2	1217	6	US-10-471-571A-4942	Sequence 4942, Ap
42	81	5.2	877	6	US-10-527-411-157	Sequence 157, App
43	81	5.2	887	6	US-10-527-411-161	Sequence 161, App
44	81	5.2	965	7	US-11-253-453-3	Sequence 3, Appl1
45	81	5.2	1842	6	US-10-511-937-2929	Sequence 2929, Ap

## ALIGNMENTS

RESULT 1  
US-10-471-571A-252  
; Sequence 252, Application US/10471571A  
; Publication No US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON Spa  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026927WO  
; CURRENT APPLICATION NUMBER: US/10/471.571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 252  
; LENGTH: 651  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)-(651)  
; OTHER INFORMATION: hypothetical ABC  
US-10-471-571A-252

Query Match	6.6%	Score 102.5;	DB 6;	Length 651;
Best Local Similarity	23.3%	Pred. No. 0.57;		
Matches	52;	Conservative	29;	Mismatches 65; Indels 77; Gaps 11;
QY	14	KEIDY-SRLPNDP-DVFR	-----CNQFYEDKYLGKCKTFTPNFFFEQY--	59
Db	112	EETDWLSKHANDYSDTYKTHMSRYESLSNQFEQLEGQVYSKIKTVLYGLNFSDFNK	171	
QY	60	-----YTLKHLIQOEYETELIMCSNVNQAHLNENFVKTFYDFPDALGVD	107	
Db	172	PINDFGGQKTRLSLAQMLLN---EPDLLLIDE-----PTNHLIDLE	209	
QY	108	FFKOLKEFNAYK-----PHEIVFNQRTSGVMCAVAIALGYKEIYLSGI-DFYQNGS	160	
Db	210	TTKWEIDLYRFKGAIVIIISHDRYFLDKIVTQIY-----DVALGDVKRVYGVNFEFIQORD	265	
QY	161	SYAFDTKQENLLKLPADFKNDRSHYIGHSKNTDIIKALE-FLEK	202	
Db	266	LYQKRMQEQY-----ESQQAIEKRLTFVEK	291	

RESULT 2  
US-10-471-571A-3348  
; Sequence 3348, Application US/10471571A

```
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3348
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(496)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-3348

Query Match          6.2%; Score 97.5; DB 6; Length 496;
Best Local Similarity 20.0%; Pred. No. 1.1; 114; Indels 95; Gaps 16;
Matches 65; Conservative 51; Mismatches 51;

Qy 12 SLKEIDYSRLPNDPDRFCNQF---YPEDKYI-----LGKKCKTVFVTPNFFF 56
Db 90 TLKHVENS---HDFRIVDQERFLVAHFQDPKYRILDVYVNHFDQSQRKVKRDFYDVRGFL 146
Qy 57 -----EQYTLKHLIQNYETELIMCSYNNQAHLENENPVKTF-----YDFPDAHLG 105
Db 147 SCSRLVDKQOTLCEFFNYPEGDTKLEKYFYKDGKPEVQKIIIVYANKQYFFANNETEIG 206
Qy 106 YDFEKQLKEFNAYKFKHEIVENQITSGVYMC-----AVATGALGYEYLSGID--- 154
Db 207 AFFIKQL-----YQHGLDFSDR---NVYTAPINLTPESIPVAVVLHSTHKNIDALD 257
Qy 155 -----FYQNGSSY---AFDTKQENLLKLAPDFKNDRSHYI-----GHSKNYDI 194
Db 258 SSPFKNVYKAMFENLSRYRAIIIVSTEQKL-----DVEKRINHPIPVNIPVGYSETIDT 312
Qy 195 KALFELBETKYK-IKLYCLCPNSLANFIELAPNLNSNFIIQEKNNYTKDILIPSE---- 249
Db 313 PVQTLQRSVKLISVARYSPKQLHQIQLIKLVS-----YVPKIELHMY 358
Qy 250 AYGFKSKNIN--FKKIKIKENVYK 272
Db 359 GFGSESKKLNELIQYGLNHHVYLR 383

RESULT 3
US-10-471-571A-1902
; Sequence 1902, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1902
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(701)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-1902
```

```
Query Match          6.1%; Score 96; DB 6; Length 701;
Best Local Similarity 19.7%; Pred. No. 2.3;
Matches 71; Conservative 46; Mismatches 92; Indels 152; Gaps 19;

Qy 15 EIDYSRLP-----NDFVFRFCNQFYFEDKYLGKKCKTVFVTPNFFF-----EQ 58
Db 324 DINFEHLPKPVVVFIDDIHDIETSTN-----YNLLNRCFEKLFKFNIGLAIAIKSTQQ 376
Qy 59 YVTLKHLI-----QNOEYETE-----LIMCSYNNQA-----HLENEN-----F 91
Db 377 FETMKQLILITFLOGNODYKTSKKLVKFMVLVFCNSMTAEIHLCHLKIKNNKKEIKYSVT 436
Qy 92 VKTFYDYFPDAHLGYDFFKQLKEFNAYF-----KHF--EIVFNQIRTS 132
Db 437 VDGFELETYSTVEQVYDVMQRLK-FHYFFIDIENSKTATHLITKNQHVHQTDTHEQ---- 491
Qy 133 GYVMCAVAIALGYEYIYL-SGIDFYQ-----NGSSYAFDTKQENLLK----- 173
Db 492 -----YKKFILDGSGISSTQFVYNNLSVSGFKYTNQDGNPQLSDIVVHLLIALL 539
Qy 174 -----LAPDFKNDRSHYIG-HSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNL 227
Db 540 RYGGGISYQLDDHSHSYISLYNKYGSPLPLMHLYKMPR-----PFVNEIDIEI 586
Qy 228 NSNFIIQEKNNYTKDILIPSEAYGKFSKNINFKIKIKENVYKLIKDLRLPSDIKH 287
Db 587 TNNYVLSKDN-----NYHF-----LLENKIND--RYMSDVYKQD 618
Qy 288 F 288
Db 619 F 619
```

```
RESULT 4
US-10-471-571A-5412
; Sequence 5412, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 5412
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-5412
```

```
Query Match          6.1%; Score 95.5; DB 6; Length 468;
Best Local Similarity 21.1%; Pred. No. 1.5;
Matches 83; Conservative 49; Mismatches 113; Indels 149; Gaps 21;

Qy 3 KVIAGNGPSLKEIDYSRLPNDPDRFCNQFYFEDKYI--GKKCKTVFVTPNFFFQYY 60
Db 109 EYVLSG---ELEKIVKHLPEKFDV-----EYGFYILFKPIPEIFVFNSSNVYKEFN 157
Qy 61 TLKHLIQNYETELIMCSYNNQAHLENE-NFVKTFYDYFPDAHLGYDFFKQLK----- 113
Db 158 T---KIESSEYRESLDSFNEILMKYLENENFINKVWKYID-----ELFKLNVQKLLI 208
Qy 114 -----EPNAVFKPHEIVFNQIRTSYVYMCVAIAIALGYEYI----- 149
Db 209 HKKIESLISSYDKKISDDCSMYSKVFEQYKKNVVLKDY-----YEVKYSNLI 259
Qy 150 -----LSGIDFYQNGSSYAFDTKQENLLKLAPDFKNDRSHYIGHS-----KNTDIKAL 197
Db 260 KVFSDVWNIMGYKKYNEY-FTRNDDTTLKIT-----NLYIGQSGFTGSGNRNVIDRIG 311
```

```

QY 198 EFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIIOEKNNYTKDIL-----IPS 247
; LENGTH: 955
; TYPE: PRT
Db 312 KGHEKLOEIMTYOPENRKTLLIFKIQPD---NIYI--TNOYKGEAVSRUKRILEOSIPN 366
; ORGANISM: Staphylococcus aureus
QY 248 SE-----AYGKFSKNIF-----KKIK--IKEN-----VYKLI 274
; SEQ ID NO 708
Db 367 KEVIDLSEISLITFFPEYKYNQHVNEFNLSLSSKIKNLAKENDGIIIIYLNFKDQNVQLI 426
; LENGTH: 955
; TYPE: PRT
QY 275 -----KDLLRLPS-DIKHYFK 289
; ORGANISM: Homo sapiens
Db 427 SEGNTFFSAKPYLHCVFENKIDILNMDKVDIENPFK 460
; SEQ ID NO 4502
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4502
; Sequence 4502, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4502
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4502

```

```

RESULT 5
US-11-293-697-4502
; Sequence 4502, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4502
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4502

Query Match 6.0%; Score 93.5; DB 7; Length 437;
Best Local Similarity 20.7%; Pred. No. 2; Mismatches 108; Indels 77; Gaps 13;
Matches 61; Conservative 48;

QY 37 DKYVLGKKCKTVYFPNPFEEQVYTLKHLIQOEYETELIMCSNYNOAHLE--NENF--- 91
; FILE REFERENCE: 2750-1590PUS2
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 48594
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(488)
; OTHER INFORMATION: Ceres Seq. ID no. 12600912
US-11-056-355B-48594

Query Match 5.7%; Score 89; DB 7; Length 488;
Best Local Similarity 20.6%; Pred. No. 5.8; Mismatches 119; Indels 138; Gaps 18;
Matches 77; Conservative 40;

QY 1 MKKVIIAGNGPSLKEIDYSLPNDPVDV-----RCNOFYFEDKYYLG-----KK 44
; LENGTH: 955
; TYPE: PRT
Db 91 MAKVMSGKPKPEVL.PV.YTELANFEGVDFWFAFASVPTSTQPNRYVHSATSHGCSNVYK 149
; ORGANISM: Staphylococcus aureus
QY 45 -----CKTVF-----YTPN-----FFEQYITLKLHQEY-----ETEL 75
; LENGTH: 955
; TYPE: PRT
Db 150 DLVKGPPQKTIPTDLSLDBENGLSFGIYQNIPTATFFKSLRLKHLVKFHSYALKFKDLAKL 209
; ORGANISM: Staphylococcus aureus
QY 76 IMCSNY-----NOAH-----LENEFVKT----- 94
; LENGTH: 955
; TYPE: PRT
Db 210 GKLPNYSVVVEQRYFDIDLFPANDDHPSHDVAAGQRFVKEVYETLRSSPQWKEMALLIYD 269
; ORGANISM: Staphylococcus aureus
QY 95 -----FYDYF-----PDHLGYDFFKQKFNAYFKFPEIYFNQRTISGVYMCVAI 141
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-11-056-355B-48594

```

```

RESULT 6
US-10-471-571A-708
; Sequence 708, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03

```

```

; SEQ ID NO 708
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-708

Query Match 5.9%; Score 92.5; DB 6; Length 955;
Best Local Similarity 20.7%; Pred. No. 6.9;
Matches 56; Conservative 44; Mismatches 72; Indels 99; Gaps 14;

QY 62 LKHLIQOEYETELIMCSNYNOAHLENEFNFKTVFYDYPDAHLGYDFFKQKFNAYFKF 121
; LENGTH: 955
; TYPE: PRT
Db 36 LKYLIRMSRTPPGLSLGSLGHFVNEP-----TRLKVGNSIQKI 76
; ORGANISM: Staphylococcus aureus
QY 122 HEIYFNQRTISGVYMCVAIAALGYKEI-YLSGID-FYQNGSSYAFDTKQENLLKLAPDFK 179
; LENGTH: 955
; TYPE: PRT
Db 77 VKV-----DGEWL-----YKLVSYIESIDEYQN-----LKVWNSK 108
; ORGANISM: Staphylococcus aureus
QY 180 -----NDRSH-----YIGHSKNTDI-----KALEFLEKTYKIKLYCLCPNSLLANFIEL 223
; LENGTH: 955
; TYPE: PRT
Db 109 AHIINDRIYLNQSAIYLNNNKDTFSIKNSSELLVFIKTT-----VTNNNI--TFSNL 159
; ORGANISM: Staphylococcus aureus
QY 224 APNLSNFIQOE-----KNNYTKDILIPSSAYGKFSKNINF--KKIKENYVYK 273
; LENGTH: 955
; TYPE: PRT
Db 160 AEKINQEFINDITKVYIHNLSVKELIYSTIRPPLSYSDNLNLYLNKLSLHNDDFVCK 219
; ORGANISM: Staphylococcus aureus
QY 274 IKDLLR-----LPSDIKHYFK 289
; LENGTH: 955
; TYPE: PRT
Db 220 IREIQKLILAYEKEIGFGEELYKDIIHMK 250
; ORGANISM: Staphylococcus aureus
US-11-056-355B-48594
; Sequence 48594, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 48594
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(488)
; OTHER INFORMATION: Ceres Seq. ID no. 12600912
US-11-056-355B-48594

```



Db 270 EHGGFYDVHTPVKGVNPDGIIGDPDF-----YFGFDRL-----GVRVPTFLISSWI 317  
Qy 142 ALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAPEKNDRSHYIG-----HSKNTDIKAL 197  
Db 318 EKGTVIHEPEGPTPHSQEHSISIPATVKKLFNLSKSHFLTKRDWAGTTEKYFRIRDSRQ 377  
Qy 198 EFLEKTYKIKLYCLCP-----NSLIANF-----IELA-----PNLNSNFIIOEK 236  
Db 378 DCPEKLPVKL-SLRPWGAKEDSKLSEFQVELIQLASQLVGDHLLNSYPDIGKNTVSEG 436  
Qy 237 NNYTKDILIPSEA 250  
Db 437 NKYAEDAVQKPLEA 450

RESULT 8  
US-11-056-355B-48593  
; Sequence 48593, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 48593  
; LENGTH: 533  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(533)  
; OTHER INFORMATION: Ceres Seq. ID no. 12600911  
US-11-056-355B-48593

Query Match 5.7%; Score 89; DB 7; Length 533;  
Best Local Similarity 20.6%; Pred. No. 6.5;  
Matches 77; Conservative 40; Mismatches 119; Indels 138; Gaps 18;  
Qy 1 MKKVIIAGNGSLKEIDYSLRPNDFDVF-----RCNQFYFDKYYLG-----KK 44  
Db 136 MAKVMSGFKPEVLVPV-YTELANEGVDFDRWFASVPTSTQPNRFYVHSATSHGCSNVKK 194  
Qy 45 -----CKTVF-----YTPN-----FPFEQYVTLKHLIQNOEY-----ETEL 75  
Db 195 DLVKGFPQKTI FDSLDEGLSFGIYYQNI PATFFKSLRLKHLVKFHSYALKPKLDAKL 254  
Qy 76 IMCSNY-----NQAH-----LENENFVKT----- 94  
Db 255 GKLPNYSVVEQRYFDIDLFPANDDHPSHDVAAGQRFVKEVYETLRSSPQWKEMALLITYD 314  
Qy 95 -----FYDYF-----PDHLGYDFFKQKLFNAYFKFHEIYFNQRI TSGVYMCVAI 141  
Db 315 EHGGFYDVHTPVKGVNPDGIIGDPDF-----YFGFDRL-----GVRVPTFLISSWI 362  
Qy 142 ALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAPEKNDRSHYIG-----HSKNTDIKAL 197  
Db 363 EKGTVIHEPEGPTPHSQEHSISIPATVKKLFNLSKSHFLTKRDWAGTTEKYFRIRDSRQ 422  
Qy 198 EFLEKTYKIKLYCLCP-----NSLIANF-----IELA-----PNLNSNFIIOEK 236  
Db 423 DCPEKLPVKL-SLRPWGAKEDSKLSEFQVELIQLASQLVGDHLLNSYPDIGKNTVSEG 481  
Qy 237 NNYTKDILIPSEA 250  
Db 482 NKYAEDAVQKPLEA 495

RESULT 9  
US-11-056-355B-48592  
; Sequence 48592, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 48592  
; LENGTH: 601  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(601)  
; OTHER INFORMATION: Ceres Seq. ID no. 12600910  
US-11-056-355B-48592

Query Match 5.7%; Score 89; DB 7; Length 601;  
Best Local Similarity 20.6%; Pred. No. 7.6;  
Matches 77; Conservative 40; Mismatches 119; Indels 138; Gaps 18;  
Qy 1 MKKVIIAGNGSLKEIDYSLRPNDFDVF-----RCNQFYFDKYYLG-----KK 44  
Db 204 MAKVMSGFKPEVLVPV-YTELANEGVDFDRWFASVPTSTQPNRFYVHSATSHGCSNVKK 262  
Qy 45 -----CKTVF-----YTPN-----FPFEQYVTLKHLIQNOEY-----ETEL 75  
Db 263 DLVKGFPQKTI FDSLDEGLSFGIYYQNI PATFFKSLRLKHLVKFHSYALKPKLDAKL 322  
Qy 76 IMCSNY-----NQAH-----LENENFVKT----- 94  
Db 323 GKLPNYSVVEQRYFDIDLFPANDDHPSHDVAAGQRFVKEVYETLRSSPQWKEMALLITYD 382  
Qy 95 -----FYDYF-----PDHLGYDFFKQKLFNAYFKFHEIYFNQRI TSGVYMCVAI 141  
Db 383 EHGGFYDVHTPVKGVNPDGIIGDPDF-----YFGFDRL-----GVRVPTFLISSWI 430  
Qy 142 ALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAPEKNDRSHYIG-----HSKNTDIKAL 197  
Db 431 EKGTVIHEPEGPTPHSQEHSISIPATVKKLFNLSKSHFLTKRDWAGTTEKYFRIRDSRQ 490  
Qy 198 EFLEKTYKIKLYCLCP-----NSLIANF-----IELA-----PNLNSNFIIOEK 236  
Db 491 DCPEKLPVKL-SLRPWGAKEDSKLSEFQVELIQLASQLVGDHLLNSYPDIGKNTVSEG 549  
Qy 237 NNYTKDILIPSEA 250  
Db 550 NKYAEDAVQKPLEA 563

RESULT 10  
US-10-953-349-10851  
; Sequence 10851, Application US/10953349  
; Publication No. US20060107345A1  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: ENCODED THERBY  
; FILE REFERENCE: 2750-1579PUS2  
; CURRENT APPLICATION NUMBER: US/10/953,349  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 40252  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 10851

```

; LENGTH: 425
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10851

Query Match
Best Local Similarity 5.6%; Score 87; DB 6; Length 425;
Matches 69; Conservative 67; Mismatches 120; Indels 96; Gaps 19;

QY 11 PSLKEIDYRLPNDVFCNCFYEDKY-----YLGKKCKTVFTYTPNFFPEQYITLKLHI 66
Db 76 PTMEETYSDVHSPAPRLNAEMEKRFKVIYIPDGDNTFYQTRKVTGKYASGYFF 135
QY 67 QN-----QEYETEL-----IMC-----SNYQAHLENENFVKTFDYDYP--DÄHL 104
Db 136 QNIRESRFTLPDPEADLFFIPISCHKMRGKGTSTYENMTVIVQNYVDGLIAPYVWNTL 195
QY 105 GYDFFKQLKFNAYKFKFHEIYNQRTISGV-----YMCVAVALGY---KEIYL-- 150
Db 196 GADHF-----FVTCDDV--GVRAFEGLPLIKTIRVCSYVNVGFIPHKDVALPQ 245
QY 151 -----SGIDFYONGSSYAFDTKQEN---LLKLAPDFKNDRSHYIGHSK-NTDIKALE 198
Db 246 VLQPPALPAGGNDVENRTTLGWAGHRNSKIRVILAHVWENDTELDISNNRINRATGLV 305
QY 199 FLEKTYKIKLYCLCP-----NS-----LLANFIELAPNLSNF-----IIQ 234
Db 306 YQKRFYRTK-FCICPGSQVNSARITDSIHYGCIPIVLSDDYDLPFDNLNWKFAVLR 364
QY 235 EKNNT-KDIL--IPSSSEAYG-----KFSKNINFKKIKIKENVYKLIKDL 277
Db 365 EODVNLKQILKNIPHSEFVSLHNNLVKQKHFQWNSPPVKDFAFHIMYEL 416

```

```

RESULT 11
US-10-953-349-10850
; Sequence 10850, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10850
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10850

```

```

Query Match
Best Local Similarity 5.6%; Score 87; DB 6; Length 429;
Matches 69; Conservative 67; Mismatches 120; Indels 96; Gaps 19;

QY 11 PSLKEIDYRLPNDVFCNCFYEDKY-----YLGKKCKTVFTYTPNFFPEQYITLKLHI 66
Db 80 PTMEETYSDVHSPAPRLNAEMEKRFKVIYIPDGDNTFYQTRKVTGKYASEGYFF 139
QY 67 QN-----QEYETEL-----IMC-----SNYQAHLENENFVKTFDYDYP--DÄHL 104
Db 140 QNIRESRFTLPDPEADLFFIPISCHKMRGKGTSTYENMTVIVQNYVDGLIAPYVWNTL 199
QY 105 GYDFFKQLKFNAYKFKFHEIYNQRTISGV-----YMCVAVALGY---KEIYL-- 150
Db 200 GADHF-----FVTCDDV--GVRAFEGLPLIKTIRVCSYVNVGFIPHKDVALPQ 249
QY 151 -----SGIDFYONGSSYAFDTKQEN---LLKLAPDFKNDRSHYIGHSK-NTDIKALE 198
Db 250 VLQPPALPAGGNDVENRTTLGWAGHRNSKIRVILAHVWENDTELDISNNRINRATGLV 309

```

```

QY 199 FLEKTYKIKLYCLCP-----NS-----LLANFIELAPNLSNF-----IIQ 234
Db 310 YQKRFYRTK-FCICPGSQVNSARITDSIHYGCIPIVLSDDYDLPFDNLNWKFAVLR 368
QY 235 EKNNT-KDIL--IPSSSEAYG-----KFSKNINFKKIKIKENVYKLIKDL 277
Db 369 EODVNLKQILKNIPHSEFVSLHNNLVKQKHFQWNSPPVKDFAFHIMYEL 420

RESULT 12
US-11-056-355B-79879
; Sequence 79879, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 79879
; LENGTH: 456
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(456)
; OTHER INFORMATION: Ceres Seq. ID no. 12653584
US-11-056-355B-79879

```

```

Query Match
Best Local Similarity 5.6%; Score 87; DB 7; Length 456;
Matches 69; Conservative 49; Mismatches 102; Indels 80; Gaps 20;

QY 22 PNDPFRNQCFYEDKY-----YLGKKCKTVFTYTPNFFPEQYITLKLHIQOEYETELI 76
Db 38 PRGFNVKSGEP-FNAPYGSYWRFMKKLCKMTKLFAG-----YQDRFVDIREET-LA 88
QY 77 MCSNYQAHLENENFVKTFDYDPAHLGYDFF-----KOLKEFNAYKFKFHE 123
Db 89 LLSTLVERSNGE-----ACDLGLEFTATTKILSKMWGKCRQNSNIPK--E 135
QY 124 IYFNQRTISGVYMCVAIAALGYKFIY--LSGIDFYQNG-----SSYAPDTKOENLLKLAP 176
Db 136 I---RKIVSDIMACAT--RFGFMELFGLRDLDFGNGKKLRSSIRWYDELVEKILK--- 187
QY 177 DKNDRSHYIGHSKNTDIKALEFLEKTY---KIKLYCLCPNLSLANFIEL-APNLSNFI 232
Db 188 EYENDKSN---EEEEKDKOIVDILLDTYNDPKAELR-LTNWQIKFFIELFMSLDTTSA 243
QY 233 IQCKNNYTKDILIPSSSEAYGKFSKNINFKKI-----KIKENYVYKL-----IKDLRL 280
Db 244 ALQ---WTWTELINHPDIFAKIRDEI--KSVVGTNTNRLIKESDLQKLPYLOAAIKETLRL 298

```

```

RESULT 13
US-11-056-355B-79878
; Sequence 79878, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR FILING DATE: 2004-02-13

```



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:26:04 ; Search time 141.657 Seconds  
(without alignments)  
8883.510 Million cell updates/sec

Title: US-10-734-719-8  
Perfect score: 876  
Sequence: 1 atgaaaaagtattattgc.....attatttcaaggagaaataa 876

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 899801 seqs, 718270062 residues

Total number of hits satisfying chosen parameters: 1799602

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New.\*  
1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
9: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	70.2	8.0	13286	6	US-10-517-441-746
C 2	67.4	7.7	5286	6	US-10-517-441-294
C 3	67.4	7.7	5286	6	US-10-517-441-568
C 4	67.4	7.7	19634	6	US-10-517-441-688
C 5	67	7.6	8900	6	US-10-517-441-428
C 6	67	7.6	13286	6	US-10-517-441-702
C 7	67	7.6	13286	6	US-10-517-441-472
C 8	64.6	7.4	5286	6	US-10-517-441-294
C 9	64.6	7.4	5286	6	US-10-517-441-568
C 10	63	7.2	10224	6	US-10-517-441-638
C 11	63	7.2	11001	6	US-10-517-441-779
C 12	62.6	7.1	19634	6	US-10-517-441-414
C 13	61.6	7.0	2501	6	US-10-517-441-562
C 14	60	6.8	8093	6	US-10-517-441-308
C 15	60	6.8	8169	6	US-10-517-441-269
C 16	59.8	6.8	10224	6	US-10-517-441-364
C 17	59.8	6.8	11001	6	US-10-517-441-505
C 18	59.4	6.8	9859	6	US-10-517-441-455
C 19	59.4	6.8	9859	6	US-10-517-441-729
C 20	59.2	6.8	6001	6	US-10-517-441-771
C 21	59	6.7	2522	7	US-11-218-305-24510
C 22	58.4	6.7	2501	6	US-10-517-441-288
C 23	58.4	6.7	8093	6	US-10-517-441-582
C 24	58.4	6.7	8169	6	US-10-517-441-543

C 25	58.4	6.7	12610	6	US-10-517-441-441	Sequence 441, App
C 26	58.2	6.6	5286	6	US-10-517-441-293	Sequence 293, App
C 27	58.2	6.6	5286	6	US-10-517-441-567	Sequence 567, App
C 28	58.2	6.6	5493	6	US-10-517-441-691	Sequence 691, App
C 29	57.4	6.6	1000	8	US-11-266-748A-281383	Sequence 281383, App
C 30	57.4	6.6	1000	8	US-11-266-748A-308023	Sequence 308023, App
C 31	57.2	6.5	12610	6	US-10-517-441-715	Sequence 715, App
C 32	57	6.5	9859	6	US-10-517-441-456	Sequence 456, App
C 33	56.8	6.5	2596	7	US-11-218-305-12130	Sequence 12130, A
C 34	56.8	6.5	4022	6	US-10-517-441-732	Sequence 732, App
C 35	56.6	6.5	5286	6	US-10-517-441-24	Sequence 24, Appl
C 36	56.6	6.5	8093	6	US-10-517-441-307	Sequence 307, App
C 37	56.6	6.5	10865	6	US-10-517-441-738	Sequence 738, App
C 38	56.4	6.4	8172	6	US-10-517-441-443	Sequence 443, App
C 39	56.4	6.4	8172	6	US-10-517-441-717	Sequence 717, App
C 40	56.2	6.4	4022	6	US-10-517-441-458	Sequence 458, App
C 41	56.2	6.4	8020	6	US-10-517-441-634	Sequence 634, App
C 42	56.2	6.4	8759	6	US-10-517-441-530	Sequence 530, App
C 43	56.2	6.4	19634	6	US-10-517-441-413	Sequence 413, App
C 44	56	6.4	2718	7	US-11-218-305-12773	Sequence 12773, A
C 45	56	6.4	6001	6	US-10-517-441-497	Sequence 497, App

## ALIGNMENTS

RESULT 1  
US-10-517-441-746/c  
; Sequence 746, Application US/10517441  
; Publication No. US20080121467A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSKENS, John  
; APPLICANT: HARBECK, Nadia  
; APPLICANT: KOENIG, Thomas  
; APPLICANT: MAIER, Sabine  
; APPLICANT: MARTENS, John  
; APPLICANT: MODEL, Fabian  
; APPLICANT: NIMMICH, Inko  
; APPLICANT: RUJAN, Tamas  
; APPLICANT: SCHMITT, Armin  
; APPLICANT: SCHMITT, Manfred  
; APPLICANT: LOOK, Maxime P.  
; APPLICANT: MARX, Almuth  
; APPLICANT: HOFFER, Heinz  
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cancer  
; TITLE OF INVENTION: proliferative disorders  
; FILE REFERENCE: 47675-93  
; CURRENT APPLICATION NUMBER: US/10/517,441  
; CURRENT FILING DATE: 2004-12-11  
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: DE 10317955.0  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: DE 10300096.8  
; PRIOR FILING DATE: 2003-01-07  
; PRIOR APPLICATION NUMBER: DE 10245779.4  
; PRIOR FILING DATE: 2002-10-01  
; NUMBER OF SEQ ID NOS: 2147  
; SEQ ID NO 746  
; LENGTH: 13286  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-746

Query Match 8.0%; Score 70.2; DB 6; Length 13286;  
Best Local Similarity 47.1%; Pred. No. 0.0011;  
Matches 317; Conservative 0; Mismatches 348; Indels 8; Gaps 3;  
QY 175 TACTACACTTTAAACATTTTAAATCCAAATATGAGACCGAATATGATGT 234  
Db 10482 TCCTTTATTACTAAAAAATCCAACTACCCCTCTAACCCCTACATATTATTCTT 10423





QY 157 AATTTCTTCTTTGAGCAATACACACTTTAAACAATTTAATCCAAAATCAAGAATATGAG 216  
Db 7372 AATTAATTTAAACAAAACAAACATAAAATAATATACCTTTAAATAATTTTTTATATAAA 7313  
QY 217 ACCGAACAAATATATGTTCTTAATACACCAAGCTCATCTAGAAAATGAAAATTTTCTGA 276  
Db 7312 ATATAAACCAACCTAAATATACAAATTAACAAAACCTTAATTAATAATTTTCCATTTTATA 7253  
QY 277 AAAATCTTTTACGAATATTTTCTGATGCTCATTTGGGATATGATTTTAAACAACAT 336  
Db 7252 AAAACCAATATATCTAATCACTTCCATATCTTAAACATATAAATTCATTTTCAATAAA 7193  
QY 337 AAAGAATTTAAGCTTATTTTAAATTCACGAATTTTCAATCAAGAATTAACCTCA 396  
Db 7192 AAAAATTTTACTTAAACACATATTTTCAATTTTCTTTTCCCATCAATATATCACTAAT 7133  
QY 397 GGGGTCTATATGTCAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGA 456  
Db 7132 TTATACCAAAAATTTTACAAAATTTTACAAAATTTTATATATTTTATATTTTATAA 7077  
QY 457 ATTGAATTTTATCAAAATGGGTCTATCTTATGCTTTTGTATACCAACAAAGAAATCTTTTA 516  
Db 7076 ACTCTATATATCTACTT--CACATCTTACTCTCAATTTTAACTTTTAACTTATCTA 7019  
QY 517 AAATAGCCCTGATTTTAAATATGATCGCTCGCATATATC--GGACATAGTAAATAATAC 575  
Db 7018 CAATTAACCTTCTAAAAAATACAAATAAATTAATAAATAAATAAATAAATAAATAAATA 6959  
QY 576 AGATATAAAGCTTTAGAAATTTCTAGAAAATTTTACAAAATTAATAAATAAATAAATA 635  
Db 6958 TCATAAAACACCCCACTCTTTCTTCAATATCAATTTTAAATAAATAAATAAATAAATA 6899  
QY 636 TCCTAATAGTCTTTTACAAAATTTTATAGAACTAGCGCAAAATTTTAAATTTCAATTTTAT 695  
Db 6898 ACACAAATACTATATAATACCCAAATTCCTCAATCTTAATAATCTTTAAATATAA 6839  
QY 696 CATACAAGAAAATAACTACATTAAGATATATCATACCTTCTAGTGGCTTTATGG 755  
Db 6838 ACTTAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6779  
QY 756 AAAATTTTCAAAAATATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 815  
Db 6778 CTTTATTAATAATTTTACCTTATAAATAATTTCTTTCTATTTATATATTTTAAATAA 6719  
QY 816 GTTGATTAAGATCTATTAAAGATTACCTAGTATATATAAGCATTA 860  
Db 6718 AAACCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6674

RESULT 5  
US-10-517-441-428/c  
; Sequence 428, Application US/10517441  
; Publication No. US20060121467A1  
; GENERAL INFORMATION:  
; APPLICANT: FOERKENS, John  
; APPLICANT: HARBECK, Nadia  
; APPLICANT: KOENIG, Thomas  
; APPLICANT: MAIER, Sabine  
; APPLICANT: MARTENS, John  
; APPLICANT: MODEL, Fabian  
; APPLICANT: NIMMICH, Inko  
; APPLICANT: RUJAN, Tamas  
; APPLICANT: SCHMITT, Armin  
; APPLICANT: SCHMITT, Manfred  
; APPLICANT: LOOK, Maxime P.  
; APPLICANT: MARX, Almuth  
; APPLICANT: HOEFELER, Heinz  
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
; TITLE OF INVENTION: proliferative disorders  
; FILE REFERENCE: 47675-93  
; CURRENT APPLICATION NUMBER: US/10/517,441  
; CURRENT FILING DATE: 2004-12-11  
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881

; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: DE 10317955.0  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: DE 10300096.8  
; PRIOR FILING DATE: 2003-01-07  
; PRIOR APPLICATION NUMBER: DE 10245779.4  
; PRIOR FILING DATE: 2002-10-01  
; NUMBER OF SEQ ID NOS: 2147  
; SEQ ID NO 428  
; LENGTH: 8900  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-428  
  
Query Match 7.6%; Score 67; DB 6; Length 8900;  
Best Local Similarity 50.4%; Pred. No. 0.0039;  
Matches 189; Conservative 0; Mismatches 185; Indels 1; Gaps 1;  
  
QY 503 AAGAAAAATCTTTTAAACCTAGCCCTGATTTTAAAAATGATCGCTCGCACTATATCGGAC 562  
Db 536 AATATATTTTATTTTAAACCATTTCTAATAATAATATATATATATATATATATATAT 477  
QY 563 ATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAACTTACAAAATAAATAC 622  
Db 476 TTACATTTTCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 417  
QY 623 TATATTTGCTTATGTC--CTAATAGTCTTTTACAAAATTTTATAGAACTAGCGCAAAATTTA 681  
Db 416 TATATATTTTATCTTCATATATATAAATTTATATATATTTTCTTAAATAATTTATCTAATTTA 357  
QY 682 AATTCAAAATTTTATCATACAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 741  
Db 356 AATTCATATTTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 297  
QY 742 AGTGGCTTATGGAAAAATTTTCAAAAAATATTAATTTTAAAAATAAATAAATAAATAAATA 801  
Db 296 AATTTAAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 237  
QY 802 AATGTTTATACAGTTGATAAAGATCTATTAAAGATTACCTAGTATATATAAGCATTAT 861  
Db 236 TACAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 177  
QY 862 TTCAAGAGAAAAATAA 876  
Db 176 TTTTAAAAATAAATAA 162

RESULT 6  
US-10-517-441-702/c  
; Sequence 702, Application US/10517441  
; Publication No. US20060121467A1  
; GENERAL INFORMATION:  
; APPLICANT: FOERKENS, John  
; APPLICANT: HARBECK, Nadia  
; APPLICANT: KOENIG, Thomas  
; APPLICANT: MAIER, Sabine  
; APPLICANT: MARTENS, John  
; APPLICANT: MODEL, Fabian  
; APPLICANT: NIMMICH, Inko  
; APPLICANT: RUJAN, Tamas  
; APPLICANT: SCHMITT, Armin  
; APPLICANT: SCHMITT, Manfred  
; APPLICANT: LOOK, Maxime P.  
; APPLICANT: MARX, Almuth  
; APPLICANT: HOEFELER, Heinz  
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
; TITLE OF INVENTION: proliferative disorders  
; FILE REFERENCE: 47675-93  
; CURRENT APPLICATION NUMBER: US/10/517,441  
; CURRENT FILING DATE: 2004-12-11  
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881





; GENERAL INFORMATION:  
 ; APPLICANT: FOEKENS, John  
 ; APPLICANT: HARBECK, Nadia  
 ; APPLICANT: KOENIG, Thomas  
 ; APPLICANT: MAIER, Sabine  
 ; APPLICANT: MARTENS, John  
 ; APPLICANT: MODEL, Fabian  
 ; APPLICANT: NIMMICH, Inko  
 ; APPLICANT: RUJAN, Tamas  
 ; APPLICANT: SCHMITT, Armin  
 ; APPLICANT: SCHMITT, Manfred  
 ; APPLICANT: LOOK, Maxime P.  
 ; APPLICANT: MARX, Almuth  
 ; APPLICANT: HOEFER, Heinz  
 ; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
 ; FILE REFERENCE: 47675-93  
 ; CURRENT APPLICATION NUMBER: US/10/517,441  
 ; CURRENT FILING DATE: 2004-12-11  
 ; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR FILING DATE: 2003-04-17  
 ; PRIOR FILING DATE: 2003-01-07  
 ; PRIOR FILING DATE: 2002-10-01  
 ; NUMBER OF SEQ ID NOS: 2147  
 ; SEQ ID NO 294  
 ; LENGTH: 5286  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-517-441-294

Query Match 7.4%; Score 64.6; DB 6; Length 5286;  
 Best Local Similarity 43.6%; Pred. No. 0.0095;  
 Matches 336; Conservative 0; Mismatches 434; Indels 1; Gaps 1;  
 QY 62 TACCAATGATTTTCATGCTATTTAGATGTAATCAATCTTTTACACCCCTAAATTTCTCTTTCGACCAACTACTA 121  
 DB 4406 TAATATAAT 4465  
 QY 122 TTGTTAAATAATGCAAGCAGTGTTCATACCCCTAAATTTCTCTTTCGACCAACTACTA 181  
 DB 4466 ATATAAT 4525  
 QY 182 CTTTAAACAATTTATCCCAAAATCAAGAAATATGAGACCGCACTAATTAATGTTCTCTAAT 241  
 DB 4526 ATATAAT 4585  
 QY 242 ACAACCAAGCTCACTAGAAAATGAAATTTTGTAAACCTTTTACGATTAATTTCTCTG 301  
 DB 4586 ATATAAT 4644  
 QY 302 ATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTAATGCTTATTTTAAAT 361  
 DB 4645 ATATAAT 4704  
 QY 362 TTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTCTATATGTTGTCAGTAGCCA 421  
 DB 4705 TAT 4764  
 QY 422 TAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTATTTTATCAAAATGGGTGAT 481  
 DB 4765 AAT 4824  
 QY 482 CTTATGCTTTTGTACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAGATG 541  
 DB 4825 ATATAAT 4884  
 QY 542 ATCGTCGCACTATATCGGACATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTAG 601

DB 4885 ATAAT 4944  
 QY 602 AAAAAAATTTACAAAATATAAATCTATATTTGCTTATGCTCTAATAGTCTCTTTAGCAAAATTTTA 661  
 DB 4945 TATATAAT 5004  
 QY 662 TAGAAGCTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAATTTTAACTTACACTA 721  
 DB 5005 TAATATAAT 5064  
 QY 722 AAGATATATCTATACCTTCTAGTGGCTTATGGAATTTTCAAAATTTTCAAAATTTTAAATTTTA 781  
 DB 5065 AAT 5124  
 QY 782 AAAAAATATAAATTTAAAGAAATTTTATTTACAAAGTTGATAAAGATCTAT 832  
 DB 5125 AT 5175

RESULT 9  
 US-10-517-441-568  
 ; Sequence 568, Application US/10517441  
 ; Publication No. US20060121467A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FOEKENS, John  
 ; APPLICANT: HARBECK, Nadia  
 ; APPLICANT: KOENIG, Thomas  
 ; APPLICANT: MAIER, Sabine  
 ; APPLICANT: MARTENS, John  
 ; APPLICANT: MODEL, Fabian  
 ; APPLICANT: NIMMICH, Inko  
 ; APPLICANT: RUJAN, Tamas  
 ; APPLICANT: SCHMITT, Armin  
 ; APPLICANT: SCHMITT, Manfred  
 ; APPLICANT: LOOK, Maxime P.  
 ; APPLICANT: MARX, Almuth  
 ; APPLICANT: HOEFER, Heinz  
 ; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
 ; FILE REFERENCE: 47675-93  
 ; CURRENT APPLICATION NUMBER: US/10/517,441  
 ; CURRENT FILING DATE: 2004-12-11  
 ; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR FILING DATE: 2003-04-17  
 ; PRIOR FILING DATE: 2003-01-07  
 ; PRIOR FILING DATE: 2002-10-01  
 ; NUMBER OF SEQ ID NOS: 2147  
 ; SEQ ID NO 568  
 ; LENGTH: 5286  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-517-441-568

Query Match 7.4%; Score 64.6; DB 6; Length 5286;  
 Best Local Similarity 43.6%; Pred. No. 0.0095;  
 Matches 336; Conservative 0; Mismatches 434; Indels 1; Gaps 1;  
 QY 62 TACCAATGATTTTCATGCTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTATC 121  
 DB 4406 TAATATAAT 4465  
 QY 122 TTGTTAAATAATGCAAGCAGTGTTCATACCCCTAAATTTCTCTTTCGACCAACTACTA 181  
 DB 4466 ATATAAT 4525  
 QY 182 CTTTAAACAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTCTATATGTTGTCAGTAGCCA 241





Db 7132 TTATACGAAAAAATTTTACGAAAAACACAATTTTACCTTTATATTTTATATTTTATAA 7077  
Qy 457 ATGATTTTATCAAAATGGGTGATCTTATGCTTTTGATACCAAAACAAGAAAATCTTTTA 516  
Db 7076 ACTCTATATATCTACTT--CAGGTCTTACTCTCTCAITTTTAAATCTTTTAACTTATCTA 7019  
Qy 517 AAACAGCCCTCATTTTAAAAATGATCGCTCGCAGCTATATC--GGACATAGTAAAAATAC 575  
Db 7018 CAATTAATCTTAAAAATACAAATAAATAAACAATAAATAAATAAATAAATAAATAAAT 6959  
Qy 576 AGATATAAAAGCTTTTGAATTTCTAGAAAAAACTTACAAAAATAAACTATATGCTTTATG 635  
Db 6958 TCATAAACAACCATTAATCTTTCTTCTATCATTTTAAATTAACAATAAATAAATAAATA 6899  
Qy 636 TCTTAATAGCTTTTATAGCAAAATTTTATAGAACTAGCGCAAAATTTAAATTTCAAAATTTAT 695  
Db 6898 ACACAAATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6839  
Qy 696 CATACAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 755  
Db 6838 ACTTAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6779  
Qy 756 AAAATTTTCAAAAAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATA 815  
Db 6778 CCTTTTATTAATTTTATTTTACCTTATATAAATTTCTTCTATTTATTTATTTTATTTT 6719  
Qy 816 GTTGATAAAGATCTATTAGATTTACCTAGTGTATGATATAAAGCAATTA 860  
Db 6718 AAACCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6674

RESULT 13  
US-10-517-441-562/c  
; Sequence 562, Application US/10517441  
; Publication No. US20060121467A1  
; GENERAL INFORMATION:  
; APPLICANT: FOEKENS, John  
; APPLICANT: HARBECK, Nadia  
; APPLICANT: KOENIG, Thomas  
; APPLICANT: MAIER, Sabine  
; APPLICANT: MARTENS, John  
; APPLICANT: MODEL, Fabian  
; APPLICANT: NIMMICH, Inko  
; APPLICANT: RUJAN, Tamas  
; APPLICANT: SCHMITT, Armin  
; APPLICANT: SCHMITT, Manfred  
; APPLICANT: LOOK, Maxime P.  
; APPLICANT: MARX, Almuth  
; APPLICANT: HOEFER, Heinz  
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
; FILE REFERENCE: 47675-93  
; CURRENT APPLICATION NUMBER: US/10/517,441  
; PRIOR FILING DATE: 2004-12-11  
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: DE 10317955.0  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: DE 10300096.8  
; PRIOR FILING DATE: 2003-01-07  
; PRIOR APPLICATION NUMBER: DE 10245779.4  
; PRIOR FILING DATE: 2002-10-01  
; NUMBER OF SEQ ID NOS: 2147  
; SEQ ID NO 562  
; LENGTH: 2501  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-562

Query Match 7.0%; Score 61.6; DB 6; Length 2501;  
Best Local Similarity 49.2%; Pred. No. 0.029;

Matches 191; Conservative 0; Mismatches 194; Indels 3; Gaps 1;  
Qy 492 TGATACCAAAACAAGAAAATCTTTTAAAACTAGCCCTCATTTTAAAAATGATCGCTCGCA 551  
Db 1744 TAATAAATAAACAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1685  
Qy 552 CTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTAGAAAAAATCTTA 611  
Db 1684 CTTTAAACAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1625  
Qy 612 CAAAATTAATACTATATGCTTTATGCTCTTAATGCTCTTAATGCTCTTTAGCAAAATTTTATAGAATAGC 671  
Db 1624 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1565  
Qy 672 GCCAAATTTTAAATTTTATCATATA--CAAGAAAAAATAAATAAATAAATAAATAAATA 728  
Db 1564 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1505  
Qy 729 ACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAAATTTTAAATAAAT 788  
Db 1504 AATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1445  
Qy 789 AAAAATTAAGAAAAATGTTTATTAACAAGTTGATTAAGATCTATTAAAGATTACCTAGTCA 848  
Db 1444 AAACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1385  
Qy 849 TATAAAGCATTTATTTTCAAGGAAAAATAA 876  
Db 1384 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1357

RESULT 14  
US-10-517-441-308/c  
; Sequence 308, Application US/10517441  
; Publication No. US20060121467A1  
; GENERAL INFORMATION:  
; APPLICANT: FOEKENS, John  
; APPLICANT: HARBECK, Nadia  
; APPLICANT: KOENIG, Thomas  
; APPLICANT: MAIER, Sabine  
; APPLICANT: MARTENS, John  
; APPLICANT: MODEL, Fabian  
; APPLICANT: NIMMICH, Inko  
; APPLICANT: RUJAN, Tamas  
; APPLICANT: SCHMITT, Armin  
; APPLICANT: SCHMITT, Manfred  
; APPLICANT: LOOK, Maxime P.  
; APPLICANT: MARX, Almuth  
; APPLICANT: HOEFER, Heinz  
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
; FILE REFERENCE: 47675-93  
; CURRENT APPLICATION NUMBER: US/10/517,441  
; PRIOR FILING DATE: 2004-12-11  
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: DE 10317955.0  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: DE 10300096.8  
; PRIOR FILING DATE: 2003-01-07  
; PRIOR APPLICATION NUMBER: DE 10245779.4  
; PRIOR FILING DATE: 2002-10-01  
; NUMBER OF SEQ ID NOS: 2147  
; SEQ ID NO 308  
; LENGTH: 8093  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-308

Query Match 6.8%; Score 60; DB 6; Length 8093;  
Best Local Similarity 45.9%; Pred. No. 0.064;

```

Matches 282; Conservative 0; Mismatches 325; Indels 7; Gaps 2;
Qy 209 AATATGACGCACTAATATGTTCTTAATACCAACCAAGCTCATCTAGAAAATGAAA 268
Db 2420 AATAACCCCTCTTAATATTTCCCTATCTTTATTCGAAAACGCACTTCGAAATACCA 2361
Qy 269 ATTGTGTAACAACTTTTACGATATTTTCCGTGATGCTATTTGGGATGATTTTTTTA 328
Db 2360 AATTTTAAATCATCTCTTAAATTTTAAACTTATAACATAACAAAAATATAAAAAA 2301
Qy 329 AACAACTTAAAGATTTAATGCTTATTTTAAATTTTCAAGAAATTTATTTCAATCAAGAA 388
Db 2300 AAAAATTTAAACATATAAAAAAATAAAAAATAAATAAATAAATAAATAAATAAATAA 2241
Qy 389 TTACTCTAGGGGTCTATATGTTGTCAGTAGCAATGCTAGGCTAGGATACAAAGAAATTTATC 448
Db 2240 ACCAAAATACACCTAAAGTTATATCAATATTTTATTAAAT-CAATTTATTAATAC 2182
Qy 449 TTTCGGGAATGATTTTATCAAAATGGGTCACTTATGCTTTTGTATACCAAAACAGAAA 508
Db 2181 GTCCGTAGCTTAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2122
Qy 509 ATCTTTTAAACTAGCCCTCAATTTTAAATGATGCTGCACTATATATCGGACATAGTA 568
Db 2121 TTATTTAAATAATATAAATAATCAAACTATCAAACTAAAAAACAAT-----AAATA 2068
Qy 569 AAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACTATAT 628
Db 2067 AAAAAACCCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2008
Qy 629 GCTTATGCTTAATAGTCTTTTAGCAATTTTATAGAACTAGCGCCAAATTTAAATCAA 688
Db 2007 CTAAAAATATTTTCTCTTTTCACTTAATAATAATAAACAATAAACAATAAACAATAA 1948
Qy 689 ATTTTATCATACAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 748
Db 1947 TTTTATTTCTACTATATTAATAATAATAAATAAATAAATAAATAAATAAATAAATAA 1888
Qy 749 CTTATGGAATTTTCAAAAAATTTAATTTTAAAAAATAAATAAATAAATAAATAAATAA 808
Db 1887 CCAAAATCTATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1828
Qy 809 ATTACAAAGTTGATA 822
Db 1827 AATAAAAAATTTTA 1814

RESULT 15
US-10-517-441-269/c
; Sequence 269, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOERKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFLE, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0

```

```

; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 269
; LENGTH: 8169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (2047, 2136, 7481, 7492, 7502, 7503, 7505, 7511)
; OTHER INFORMATION: unknown base
US-10-517-441-269

Query Match 6.8%; Score 60; DB 6; Length 8169;
Best Local Similarity 45.6%; Pred. No. 0.065;
Matches 331; Conservative 0; Mismatches 385; Indels 10; Gaps 3;

Qy 97 TTTTATTTTGAAGATAAATACTATCTTGGTAAATAATGCAAGCAGTGTTTTACACCCCT 156
Db 6621 TCTTATTTTATTAATAAACAATAATTTTACTATATTTATCCAAACTAATCTTTAAATCTCTAT 6562
Qy 157 AATTTCTCTTTGAGCAA-----TACTACACTTTTAAAAACATTTTAAATCCAAAAATCAAG 208
Db 6561 ATTCAATAATATCTCTCTACATCAACTCCCAAAATATTAATAATACAAATATACACACAC 6502
Qy 209 AATATGAGACCGAATTAATTTATGTTCTTAATTTACAAACCAAGCTCATCTAGAAAAATGAAA 268
Db 6501 ATACCCACCCCATATTAATAATTTTCATTTTAAAAAATAAATAATTTTACTAATTTATAT 6442
Qy 269 ATTTTGTAAAAAATTTTACGATTTTTCCTGATGCTCATTTGGGATATGATTTTTTTA 328
Db 6441 ATCTTCAAACTTAACTTACTTCTTAATTTTAAATCTTAACTTCTTTTAAAAAATAATAATTTCTA 6382
Qy 329 AACAACTTAAAGAAATTTAAT-GCTTATTTTAAATTTTCAAGAAATTTTATTTCAATCAAGA 387
Db 6381 AAATATTTTAACTTAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6322
Qy 388 ATTACTCTAGGGGTCTATATGTTGTCAGTAGGCATAGCCCTAGGATACAAAGAAATTTAT 447
Db 6321 AAAACCAAAAAAATTAATAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6262
Qy 448 CTTTCGGGAATTTGATTTTATCAAAAT-GGGTCACTTTATGCTTTTGTATACCAACAAAGA 506
Db 6261 ATCTAAACCATTCATTTTACAACTTTCAAAATAAATACTTAAACACTTTTATTTCCAACTATA 6202
Qy 507 AAATCTTTTAAACTAGCCCTGATTTTAAATAATGATCGCTCGCACTATATCGGACATAG 566
Db 6201 AAATATAATTAATTAATCATCATCTAATTAATAATAAATAAATAAATAAATAAATAAATAAATAA 6142
Qy 567 TAAAAATACAGATATAAAGCTTTTAGAATTTCTAGAAAAAATTTTACAAAAATAAATAAATAA 626
Db 6141 ACACAATTTCTTTTAAACAATAAACAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 6082
Qy 627 TTGCTTATGCTCTAATAGTCTTTTACGAAATTTTATAGAACTAGCGCCAAATTTAAATTC 686
Db 6081 TCTACTACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6022
Qy 687 AAATTTTATACAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 746
Db 6021 TTCTTAAATCTTCAGCACTAAAAATTTCTTAATACAAAAATCTTAACCTTATCTTACA 5962
Qy 747 GCGTTATGAAAAATTTTCAAAAAATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 806
Db 5961 CTCTCCAACTATTTATAAATCCATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5902
Qy 807 TTATTA 812
Db 5901 CTCTTA 5896

```

Search completed: July 19, 2006, 15:44:56  
Job time : 142.657 secs

---

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:53:32 ; Search time 655.162 Seconds  
(without alignments)  
9322.405 Million cell updates/sec

Title: US-10-734-719-8  
Perfect score: 876  
Sequence: 1 atgaaaaagtattattgc.....attatttcaaggaataaa 876

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_8.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*
- 15: Geneseqn2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	876	100.0	876	3	AAA53726
2	876	100.0	876	8	ABT13669
3	869.8	99.3	873	8	ABT13671
4	866.6	98.9	873	8	ABT13670
5	844	96.3	876	3	AAA53724
6	844	96.3	876	8	ABT13667
7	842.4	96.2	876	13	ADU77566
8	842.4	96.2	12219	13	ADU77595
9	837.6	95.6	11474	3	AAA53720
10	837.6	95.6	11474	8	ABT13665
11	836	95.4	876	3	AAA53721
12	836	95.4	876	8	ABT13666
13	820	93.6	876	3	AAA53725
14	820	93.6	876	8	ABT13668
15	311.6	35.6	1293	2	AAZ225693
16	286.6	32.7	110000	2	AAZ20693_03
17	281.8	32.2	191996	13	ADT05647
18	265.4	30.3	13379	13	ADT05530

C 19	265.4	30.3	117576	13	ADT05736	Adt05736 Haemophil
C 20	265.4	30.3	349980	13	ADT05648	Adt05648 Haemophil
C 21	260.4	29.7	4277	13	ADT05421	Adt05421 Haemophil
C 22	255	29.1	909	8	ACC71694	Acc71694 alpha-2,3
C 23	90.2	10.3	2482	14	ADZ71091	Adz71091 Human chr
C 24	89	10.2	8056	8	ABZ10246	Abz10246 Haematopo
C 25	88.8	10.1	8056	8	ABZ10246	Abz10246 Haematopo
C 26	86.2	9.8	19087	6	ABL32793	Abz132793 Human inm
C 27	82.6	9.4	8056	8	ABZ10100	Abz10100 Haematopo
C 28	81.4	9.3	158001	12	ADL17884	Adl17884 Human pho
C 29	80.4	9.2	15548	6	ABL34155	Abi34155 Human inm
C 30	79.6	9.1	1554	14	ADZ70999	Adz70999 Human chr
C 31	78	8.9	8056	8	ABZ10100	Abz10100 Haematopo
C 32	76.6	8.7	6465	6	ABL32985	Abi32985 Human inm
C 33	76	8.7	17131	6	ABL33053	Abi33053 Human inm
C 34	75.6	8.6	1501	8	ABZ10188	Abz10188 Haematopo
C 35	75.6	8.6	1501	10	ADE84162	Ade84162 Human lym
C 36	74	8.4	18988	4	AAS46342	Aas46342 Tumour su
C 37	74	8.4	18988	6	ABL32701	Abi32701 Human inm
C 38	74	8.4	18988	6	ABL34509	Abi34509 Human met
C 39	74	8.4	18988	6	ABL70204	Abi70204 Chemical
C 40	74	8.4	18988	7	ADZ99770	Adz99770 Complemen
C 41	73.8	8.4	5930	6	ABL32517	Abi32517 Human inm
C 42	73.4	8.4	12592	6	AAS61102	Aas61102 Human gen
C 43	73	8.3	110000	13	ABD32968_6	Continuation (7 of
C 44	72.6	8.3	5689	4	AAS45384	Aas45384 Chemical
C 45	72.6	8.3	5689	4	AAS46426	Aas46426 Tumour su

## ALIGNMENTS

### RESULT 1

AAA53726	
ID	AAA53726 standard; DNA; 876 BP.
XX	
AC	AAA53726;
XX	
DT	15-SRP-2003 (revised)
DT	22-DEC-2000 (first entry)
XX	
DE	Campylobacter jejuni O:19 serotype CstII sialyltransferase.
XX	

Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
sialic acid synthase; CMP-sialic acid synthetase; mimetic; antibody;  
immunity; immunogen; ganglioside; ds.

Campylobacter jejuni; O:19 serotype.

PH	Key	Location/Qualifiers
FT	CDS	1..876
FT		/*tag= a
FT		/product= "CstII sialyltransferase"

WO2000046379-A1.

10-AUG-2000.

01-FEB-2000; 2000WO-CA000086.

01-FEB-1999; 99US-0118213P.

31-JAN-2000; 2000US-00495406.

(CANADA ) NAT RES COUNCIL CANADA.

Gilbert M, Wakarchuk WW;

WPI; 2000-524418/47.

P-PSDB; AAY97212.

Novel glycosyltransferase polypeptides and polynucleotides useful for

PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 XX reagents and as immunogen for producing antibodies.  
 PS Disclosure; Page 96; 120pp; English.  
 XX  
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesizing sialylated oligosaccharide such as ganglioside,  
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyze the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 15-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Db 541 GATCGCTCCGACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
 QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660  
 Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTATAGTCTTTTAGCAAAATTTT 660  
 QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATACTACACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATACTACACT 720  
 QY 721 AAGATATATCTATCATCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATTAATTAATTTT 780  
 Db 721 AAGATATATCTATCATCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATTAATTAATTTT 780  
 QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAGATCTATTAAAGATTA 840  
 Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAGATCTATTAAAGATTA 840  
 QY 841 CCTAGTGATATATAAGCATTATTTCAAAGGAAAAATAA 876  
 Db 841 CCTAGTGATATATAAGCATTATTTCAAAGGAAAAATAA 876  
 RESULT 2  
 ABT13669  
 ID ABT13669 standard; DNA; 876 BP.  
 XX  
 AC ABT13669;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE C. jejuni bifunctional sialtransferase cstII coding sequence #4.  
 XX  
 KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetyl-galactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 XX  
 PN WO200274942-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 22-FEB-2002; 2002WO-CA000229.  
 XX  
 PR 21-MAR-2001; 2001US-00816028.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Gilbert M, Wakarchuk WW;  
 XX  
 DR WPI; 2003-040554/03.  
 DR P-PSDB; ABJ18482.  
 XX  
 PT New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Disclosure; Page 98; 107pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C.jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 CC Acetyl-galactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune



CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention  
 XX

SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Query Match 100.0%; Score 876; DB 8; Length 876;

Best Local Similarity 100.0%; Pred. No. 4.3e-134;

Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTATTATTCGTGAAATGGACCAAGTTTAAAGAAATGATTATTCACAGG 60

DB 1 ATGAAAAAGTATTATTCGTGAAATGGACCAAGTTTAAAGAAATGATTATTCACAGG 60

QY 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

DB 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATTTATGTTCTTAAT 240

DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATTTATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

DB 241 TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300

QY 301 GATGCTCATTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360

DB 301 GATGCTCATTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA 360

QY 361 TTTTCAGAAATTTTTCATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420

DB 361 TTTTCAGAAATTTTTCATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGTCA 480

DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540

DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTGCACATATCGGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600

DB 541 GATCGCTGCACATATCGGACATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTACAAATATAATCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 660

DB 601 GAAAAAATTTACAAATATAATCTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCAACAGAAAAATACTACACT 720

DB 661 ATAGAACTAGCGCCAAATTTAAATTTTATCAACAGAAAAATACTACACT 720

QY 721 AAAGATATATCTATACCTTTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780

DB 721 AAAGATATATCTATACCTTTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780

QY 781 AAAAAAATAAATAAAGAAATCTTTATTAACAGTTTATTAAGATCTATTAAGATTA 840

DB 781 AAAAAAATAAATAAAGAAATCTTTATTAACAGTTTATTAAGATCTATTAAGATTA 840

QY 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAAATAA 876

RESULT 3

ABT13671

ID ABT13671 standard; DNA; 873 BP.

XX AC ABT13671;

XX XX 07-FEB-2003 (first entry)

XX C. jejuni bifunctional sialtransferase cstII coding sequence #6.

XX Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 XX GalNAC transferase; N-Acetylgalactosamine transferase;  
 XX galactosyltransferase; sialyltransferase; sialic acid synthase;  
 XX cytidine 5'-monophosphate sialic acid synthetase;  
 XX CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 XX ganglioside mimetics; inflammation; tumour metastasis.

XX Campylobacter jejuni.

XX WO200274942-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-CA000229.

XX 21-MAR-2001; 2001US-00816028.

XX (CANA ) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2003-040554/03.

XX P-PSDB; ABJ18485.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 XX mechanisms of organisms that synthesize ganglioside mimetics.

XX Disclosure; Page 99; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention

SQ Sequence 873 BP; 353 A; 117 C; 107 G; 296 T; 0 U; 0 Other;

Query Match 99.3%; Score 869.8; DB 8; Length 873;

Best Local Similarity 99.8%; Pred. No. 4.4e-133;

Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTATTATTCGTGAAATGGACCAAGTTTAAAGAAATGATTATTCACAGG 60

DB 1 ATGAAAAAGTATTATTCGTGAAATGGACCAAGTTTAAAGAAATGATTATTCACAGG 60

QY 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

DB 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

DB 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGCACTAATTTATGTTCTTAAT 240



Db 541 GATCGCTCACATATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAATAAAATTAATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAATAAAATTAATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAAATACTACACT 720  
Db 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAAATACTACACT 720  
QY 721 AAAGATATATCTATACCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 721 AAAGATATATCTATACCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAAATTTAAATTAAGAAAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATTTAAATTAAGAAAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 5  
AAAS3724  
ID AAAS3724 standard; DNA; 876 BP.  
XX AAAS3724;  
XX AC  
XX AC  
DT 15-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.  
XX  
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside; ds.  
XX  
OS Campylobacter jejuni; O:10 serotype.  
XX  
FH Key Location/Qualifiers  
FT 1. .876  
FT /\*tag= a  
FT /product= "alpha-2,3-sialyltransferase"  
XX  
PN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
XX 01-FEB-2000; 2000WO-CA000086.  
PF  
XX 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
PI Gilbert M, Wakarchuk WW;  
XX WPI; 2000-524418/47.  
DR P-PSDB; AAY97210.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
XX Claim 6; Page 92; 120pp; English.  
PS  
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesizing sialylated oligosaccharide such as ganglioside,  
CC lysoganglioside or their mimics. Glycosyltransferases are useful for

CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyze the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 15-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;  
Query Match 96.3%; Score 844; DB 3; Length 876;  
Best Local Similarity 97.7%; Pred. No. 7.1e-129;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
QY 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120.  
Db 61 CTACCAATGATTTTGATGTATTAGATGCAATTTTATTTTGAAGATAAATACTAT 120  
QY 121 CTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATACTAC 180  
Db 121 CTGGTAAAAAATGCAAAAGCAGTGTTTTACAAATCTCTGCTCTTTTGTGAACAATACTAC 180  
QY 181 ACTTTAAACAATTTAATCCAAATCAAGATATGAGACCGCACTAATTTATGTGTTCTAAT 240  
Db 181 ACTTTAAACAATTTAATCCAAATCAAGATATGAGACCGCACTAATTTATGTGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAA 360  
QY 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTTACGAAATTTTATCTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCT 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAAATGGGCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAAATGGGCA 480  
QY 481 TCTTATGCTTTTGTATACCAAGAAATTTTAAAAAATCTAGCCCTGATTTTAAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAAGAAATTTTAAAAAATCTAGCCCTGATTTTAAAAAAT 540  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTTACAAAAATAAATACTATTTGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTTACAAAAATAAATACTATTTGCTTATGCTTAACTATTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAAATACTACACT 720

Db 661 ATAGAAGTACGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAAATACATACACT 720  
Qy 721 AAAGATATACATACACTTCTAGAGAGCTTATGGAATAATTTTCAAAAAATTTAAATTT 780  
Db 721 AAAGATATACATACACTTCTAGAGAGCTTATGGAATAATTTTCAAAAAATTTAAATTT 780  
Qy 781 AAAAAATATAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGTATATAAGCATTTATTTCAAGGAAAAATAA 876  
Db 841 CCTAGTGTATATAAGCATTTATTTCAAGGAAAAATAA 876

## RESULT 6

ABT13667

ID ABT13667 standard; DNA; 876 BP.

XX AC ABT13667;

XX DT 07-FEB-2003 (first entry)

XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #2.

XX Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;

KW GalNac transferase; N-Acetylgalactosamine transferase;

KW galactosyltransferase; sialyltransferase; sialic acid synthase;

KW cytidine 5'-monophosphate sialic acid synthetase;

KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;

KW ganglioside mimetics; inflammation; tumour metastasis.

XX Campylobacter jejuni.

XX W0200274942-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-CA000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX PA (CANADA) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX DR WPI; 2003-040554/03.

XX DR P-PSDB; ABJ18480.

XX PT New glycosyltransferases from Campylobacter, useful for synthesizing

XX PT gangliosides and ganglioside mimetics, and in studying the pathogenesis

XX PT mechanisms of organisms that synthesize ganglioside mimetics.

XX PS Claim 8; Page 96-97; 107pp; English.

XX The invention comprises the amino acid and coding sequences of

XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention

XX may be either an: acyltransferase; glycosyltransferase; GalNac (N-

XX Acetylgalactosamine) transferase; galactosyltransferase;

XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)

XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein

XX sequences of the invention are useful for ganglioside synthesis, studying

XX ganglioside mimetics, and for designing oligonucleotides to inhibit

XX expression of Campylobacter enzymes involved in the biosynthesis of

XX ganglioside mimetics that can mask the pathogen's from the host's immune

XX system. The C. jejuni oligosaccharides of the invention may be used as

XX diagnostic reagents (e.g. to locate areas of inflammation or tumour

XX metastasis). The present DNA sequence represents a Campylobacter jejuni

XX gene of the invention

XX SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

XX Query Match 96.3%; Score 844; DB 8; Length 876;

Best Local Similarity 97.7%; Pred. No. 7.le-129;  
Matches 856; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTTATTATTGCTCGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Db 1 ATGAAAAAGTTTATTATTGCTCGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 61 CTACCAAAATGATTTTGTATGTTTATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120

Db 61 CTACCAAAATGATTTTGTATGTTTATGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180

Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180

Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAATCAATATATGCTGTTCTAAT 240

Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAATCAATATATGCTGTTCTAAT 240

Qy 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCCCT 300

Db 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCCCT 300

Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360

Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360

Qy 361 TTTACGAGAAATTTTATTTCAATCAAGAAATTTACTCTAGGGGTCTATATGCTGAGTAGCC 420

Db 361 TTTACGAGAAATTTTATTTCAATCAAGAAATTTACTCTAGGGGTCTATATGCTGAGTAGCT 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTGGCTCTGATTTTAAAAAT 540

Db 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATCTTTTAAAACTGGCTCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

Qy 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAAATAGTCTTTTAGCAAAATTTT 660

Db 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAAATAGTCTTTTAGCAAAATTTT 660

Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAGAAAAATAAATACTACACT 720

Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAGAAAAATAAATACTACACT 720

Qy 721 AAAGATATACATACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780

Db 721 AAAGATATACATACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780

Qy 781 AAAAAATAAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Db 781 AAAAAATAAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGTATAAAGCATTTATTTCAAGGAAAAATAA 876

Db 841 CCTAGTGTATAAAGCATTTATTTCAAGGAAAAATAA 876

## RESULT 7

ADU77566

ID ADU77566 standard; DNA; 876 BP.

XX AC ADU77566;

XX DT 10-FEB-2005 (first entry)

XX DE Sialyltransferase (cstII).

XX gene amplification; campylobacter infection; Sialytransferase; cstII; ds.  
KW Campylobacter jejuni.  
XX WO2004101817-A1.  
XX 25-NOV-2004.  
XX 19-MAY-2004; 2004WO-AU000654.  
XX 19-MAY-2003; 2003AU-00902445.  
XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
XX Shi F, Fry BN, Coloe PJ;  
XX WPI; 2004-834013/82.  
XX Identifying Campylobacter in a sample to comprises comparing the profile  
PT of an amplification product to a comparator Campylobacter species after  
PT performing nucleic acid amplification of the wla gene cluster.  
XX  
XX Claim 6; SEQ ID NO 13; 75pp; English.  
XX  
XX The invention describes a method of identifying Campylobacter in a sample  
CC comprising comparing the profile of an amplification product to a  
CC comparator Campylobacter species after performing nucleic acid  
CC amplification. Identifying Campylobacter in a sample comprises: (a)  
CC performing nucleic acid amplification by contacting the sample with a  
CC pair of nucleic acid primers, where at least one nucleic acid primer is  
CC derived from the wla gene cluster of Campylobacter, for a time and under  
CC conditions for generation of an amplification product comprising the wla  
CC gene cluster of Campylobacter or its portion; (b) preparing a profile of  
CC the amplification product; and (c) comparing the profile to a profile of  
CC a comparator Campylobacter species, strain or type, and thus, determining  
CC the identity of Campylobacter in the sample. Also described are: (a) a  
CC pair of nucleic acid primers, where at least one nucleic acid primer is  
CC derived from the wla gene cluster of Campylobacter; or (b) a pair of  
CC nucleic acid primers for amplification of the wla gene cluster of  
CC Campylobacter or its portion; or (c) a probe capable of hybridizing to  
CC the wla gene cluster of Campylobacter. The method and kit are useful for  
CC identifying and detecting the presence of Campylobacter in a sample, or  
CC for diagnosing infections associated with Campylobacter. This sequence  
CC represents Sialytransferase, encoded by a gene in the Campylobacter  
CC jejuni wla cluster used in the inventive method of detecting  
CC Campylobacter jejuni infection.  
XX  
XX Sequence 876 BP; 356 A; 115 C; 104 G; 301 T; 0 U; 0 Other;  
SQ

Query Match 96.2%; Score 842.4; DB 13; Length 876;  
Best Local Similarity 97.6%; Pred. No. 1.3e-128;  
Matches 855; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
DB 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
QY 61 CTACCAATGATTTGATGTTATTTAGATGATATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTGATGTTATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTCTTGACGAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTATTTTACAACTCTGCTCTTTTGTGAACAATACTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGACCGGAATAATTTATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAATATGACCGGAATAATTTATGTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTCT 300

QY 301 GATGCTCATTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAA 360  
DB 301 GATGCTCATTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAA 360  
QY 361 TTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGTCTATATGTCGACGTAGCC 420  
DB 361 TTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGTCTATATGTCGACGTAGCT 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTATGATACCAAAATCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTATGATACCAAAATCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAACTTACAAAAATAAAACTATATTTGCTTATGCTCTAATAGTCTTTTACAAATTTT 660  
DB 601 GAAAAAACTTACAAAAATAAAACTATATTTGCTTATGCTCTAATAGTCTTTTACAAATTTT 660  
QY 661 ATAGAACTAGCGCAAAATTTTAAATTTTATCATACAGAAAAATAAATACTACACT 720  
DB 661 ATAGAACTAGCGCAAAATTTTAAATTTTATCATACAGAAAAATAAATACTACACT 720  
QY 721 AAGATATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTATTTT 780  
DB 721 AAGATATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTATTTT 780  
QY 781 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAGATTA 840  
DB 781 AAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAGATTA 840  
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAAATAA 876

RESULT 8  
ADU77595  
ID ADU77595 standard; DNA; 12219 BP.  
XX  
AC ADU77595;  
XX  
DT 10-FEB-2005 (first entry)  
XX  
DE Campylobacter jejuni strain ATCC 43469 wla cluster.  
XX  
KW gene amplification; campylobacter infection; wla; ds.  
XX  
OS Campylobacter jejuni.  
XX WO2004101817-A1.  
XX  
PD 25-NOV-2004.  
XX  
PF 19-MAY-2004; 2004WO-AU000654.  
XX  
PR 19-MAY-2003; 2003AU-00902445.  
XX  
PA (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
XX  
PI Shi F, Fry BN, Coloe PJ;  
XX  
DR WPI; 2004-834013/82.  
XX  
PT Identifying Campylobacter in a sample to comprises comparing the profile  
PT of an amplification product to a comparator Campylobacter species after  
PT performing nucleic acid amplification of the wla gene cluster.



FT CDS /note= "Open reading frame 10a"  
 complement(10554..11366)  
 /tag= j  
 /product= "Lipooligosaccharide biosynthetic enzyme"  
 /note= "Open reading frame 12a"  
 WO200046379-A1.  
 XX 10-AUG-2000.  
 XX 01-FEB-2000; 2000WO-CA000086.  
 XX 01-FEB-1999; 99US-0118213P.  
 PR 31-JAN-2000; 2000US-00495406.  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 PA Gilbert M, Wakarchuk WW;  
 PI WPI; 2000-524418/47.  
 DR P-PSDB; AAY97200, AAY97201, AAY97202, AAY97203, AAY97204, AAY97205,  
 DR AAY97206.  
 XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 PT reagents and as immunogen for producing antibodies.  
 XX Claim 1; Page 86-90; 120pp; English.  
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesizing sialylated oligosaccharides such as ganglioside,  
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyze the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 15-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;  
 Query Match 95.6%; Score 837.6; DB 3; Length 11474;  
 Best Local Similarity 97.3%; Pred. No. 7.7e-128;  
 Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 DB 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAG 6107  
 QY 61 CTACCAATGATTTTGATGCTATTATAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120  
 DB 6108 CTACCAATGATTTTGATGCTATTATAGATGTAATCAATTTTATTTGAAGATAAATCTAT 6167  
 QY 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
 DB 6168 CTGTGTAATAAATGCAAGCAGTGTTTTACATCTCTTTCTTTTGGACAACTACTAC 6227  
 QY 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAACCTAATTTATGTTCTAAT 240  
 |||||||

Db 6228 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAATTAATTTATGTTCTAAT 6287  
 QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAACCTTTTACGATTATTTTCCT 300  
 |||||||  
 Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAATTTTGTAAACCTTTTACGATTATTTTCCT 6347  
 |||||||  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
 |||||||  
 Db 6348 GATGCTCATTTGGGATATGATTTTCAACAACTTAAAGAAATTTTAAATGCTTATTTTAA 6407  
 |||||||  
 QY 361 TTTCAACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGAGTAGCC 420  
 |||||||  
 Db 6408 TTTCAACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGAGTAGCC 6467  
 |||||||  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 |||||||  
 Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
 |||||||  
 QY 481 TCTTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 |||||||  
 Db 6528 TCTTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 6587  
 |||||||  
 QY 541 GATCGCTCGCATATATCGGACATACATAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600  
 |||||||  
 Db 6588 GATAATTCACACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 6647  
 |||||||  
 QY 601 GAAAAAACTTACAAAATAAAACTATATTTGCTTCTTAATAGTCTTTTACAAATTTT 660  
 |||||||  
 Db 6648 GAAAAAACTTACAAAATAAAACTATATTTGCTTCTTAATAGTCTTTTACAAATTTT 6707  
 |||||||  
 QY 661 ATAGAACTAGCGCAATTTTAAATTTTAAATTTTATCATACAGAAAAAATAACCTACACT 720  
 |||||||  
 Db 6708 ATAGAACTAGCGCAATTTTAAATTTTAAATTTTATCATACAGAAAAAATAACCTACACT 6767  
 |||||||  
 QY 721 AAAGATATACATACCTCTAGTGAGCTTATCGAAAAATTTTCAAAAAATTTTAAATTTT 780  
 |||||||  
 Db 6768 AAAGATATACATACCTCTAGTGAGCTTATCGAAAAATTTTCAAAAAATTTTAAATTTT 6827  
 |||||||  
 QY 781 AAAAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAGATTA 840  
 |||||||  
 Db 6828 AAAAAATAAAAAATTAAGAAAAATTTTATTTACAAAGTTGATAAAGATCTATTAAGATTA 6887  
 |||||||  
 QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876  
 |||||||  
 Db 6888 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 6923  
 |||||||  
 RESULT 10  
 ABL13665  
 ID ABL13665 standard; DNA; 11474 BP.  
 XX  
 AC ABL13665;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE Campylobacter jejuni genomic lipooligosaccharide biosynthesis locus.  
 XX  
 KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylglucosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 OS  
 XX WO200274942-A2.  
 PN  
 XX 26-SEP-2002.  
 PD  
 XX 22-FEB-2002; 2002WO-CA0000229.  
 PF  
 XX 21-MAR-2001; 2001US-00816028.  
 PR



XX (CANA ) NAT RES COUNCIL CANADA.  
 XX Gilbert M, Wakarchuk WW;  
 XX WPI; 2003-040554/03.  
 XX  
 XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 XX mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 XX Claim 1; Page 89-95; 107pp; English.  
 XX  
 XX The invention comprises the amino acid and coding sequences of  
 XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 XX may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 XX Acetylglucosamine) transferase; galactosyltransferase;  
 XX sialyltransferase; sialic acid synthetase; cytidine 5'-monophosphate (CMP)  
 XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 XX sequences of the invention are useful for ganglioside synthesis, studying  
 XX ganglioside mimetics, and for designing oligonucleotides to inhibit  
 XX expression of Campylobacter enzymes involved in the biosynthesis of  
 XX ganglioside mimetics that can mask the pathogen's from the host's immune  
 XX system. The C. jejuni oligosaccharides of the invention may be used as  
 XX diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 XX metastasis). The present DNA sequence represents a Campylobacter jejuni  
 XX gene of the invention  
 XX  
 XX Query Match 95.6%; Score 837.6; DB 8; Length 11474;  
 XX Best Local Similarity 97.3%; Pred. No. 7.7e-128;  
 XX Matches 852; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTATTATGCTGGAAATGCGAACAGTAAAGAAATGATTATTCAGG 60  
 Db 6048 ATGAAAAAGTATTATGCTGGAAATGCGAACAGTAAAGAAATGATTATTCAGG 6107  
 Qy 61 CTACCAATGATTTGATGATTTAGATGATCAATTTATTTTGAAGATAATACATAT 120  
 Db 6108 CTACCAATGATTTGATGATTTAGATGATCAATTTATTTTGAAGATAATACATAT 6167  
 Qy 121 CTGCTGAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACATAC 180  
 Db 6168 CTGCTGAAAAATGCAAGCAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACATAC 6227  
 Qy 181 ACTTAAACATTTAATCAAAATCAAGATATGAGACCGAAGTAAATGATGTTCTAAT 240  
 Db 6228 ACTTAAACATTTAATCAAAATCAAGATATGAGACCGAAGTAAATGATGTTCTAAT 6287  
 Qy 241 TACACCAAGCTCATCTAGAAATCAAAATTTGTAAAACTTTTACGATTTATTTCTCT 300  
 Db 6288 TACACCAAGCTCATCTAGAAATCAAAATTTGTAAAACTTTTACGATTTATTTCTCT 6347  
 Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360  
 Db 6348 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 6407  
 Qy 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCAGTAGCC 420  
 Db 6408 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCAGTAGCC 6467  
 Qy 421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTC 480  
 Db 6468 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTC 6527  
 Qy 481 TCTTATGCTTTTGTATACCAAGAAATTTTAAAACTAGCCCTGATTTTAAAAAT 540  
 Db 6528 TCTTATGCTTTTGTATACCAAGAAATTTTAAAACTAGCCCTGATTTTAAAAAT 6587  
 Qy 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCT 600  
 Db 6588 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCT 6647

Qy 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 660  
 Db 6648 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAAAATTTT 6707  
 Qy 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAATACTACACT 720  
 Db 6708 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAATACTACACT 6767  
 Qy 721 AAAGATATACCTACCTCTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
 Db 6768 AAAGATATACCTACCTCTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 6827  
 Qy 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAACTTGATAAAAGATCTTATTAAAGATTA 840  
 Db 6828 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAACTTGATAAAAGATCTTATTAAAGATTA 6887  
 Qy 841 CTTAGTATATTAAGCAATTTTCAAGGAAAAATAA 876  
 Db 6888 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 6923  
 XX  
 XX RESULT 11  
 XX AAA53721  
 XX ID AAA53721 standard; DNA; 876 BP.  
 XX AC AAA53721;  
 XX AC  
 XX 15-SEP-2003 (revised)  
 XX 22-DEC-2000 (first entry)  
 XX  
 XX Campylobacter jejuni OH4384 CstII sialyltransferase coding sequence.  
 XX  
 XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
 XX acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
 XX Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
 XX sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
 XX immunity; immunogen; ganglioside; ds.  
 XX  
 XX Campylobacter jejuni; OH4384.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 1..876  
 XX FT /\*tag= a  
 XX FT /product= "CstII sialyltransferase"  
 XX  
 XX WO200046379-A1.  
 XX  
 XX 10-AUG-2000.  
 XX  
 XX 01-FEB-2000; 2000WO-CA000086.  
 XX  
 XX 01-FEB-1999; 99US-0118213P.  
 XX 31-JAN-2000; 2000US-00495406.  
 XX  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 XX Gilbert M, Wakarchuk WW;  
 XX WPI; 2000-524418/47.  
 XX P-PSDB; AAY97204.  
 XX  
 XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
 XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 XX reagents and as immunogen for producing antibodies.  
 XX  
 XX Claim 6; Page 90; 120pp; English.  
 XX  
 XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 XX useful for synthesizing sialylated oligosaccharides such as ganglioside,  
 XX lysoganglioside or their mimics. Glycosyltransferases are useful for  
 XX chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 XX other oligosaccharides that have biological activity. The enzymes and



551 ATAGTACIAGCGCCAAATTTTAAATTCAAATTTTATCATACAAAGAAAAATAACTACACT 720

### Matches

```

Query Match          95.4%; Score 836; DB 8; Length 876;
Best Local Similarity 97.1%; Pred. No. 1.4e-127;
Matches 851: Conservative 0; Mismatches 25; Indels 0; Gaps 0

```

QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
DB |||||  
1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 60  
QY 61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120  
DB |||||  
61 CTACCAATGATTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120  
QY 121 CTTGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTGTGAGCAATACTAC 180  
DB |||||  
121 CTTGTAAAAATGCAAGCAGTGTTTTACAAATCTCTATTTCTTTTTTTTGAACAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240  
DB |||||  
181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATTTATTTCT 300  
DB |||||  
241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATTTATTTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAA 360  
DB |||||  
301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAA 360  
QY 361 TTTTCAAGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
DB |||||  
361 TTTTCAAGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAAATGGGTCA 480  
DB |||||  
421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAT 540  
DB |||||  
481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAT 540  
QY 541 GATCGCTCGCACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
DB |||||  
541 GATTAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
QY 601 GAAAAAATTTCAAAAATTTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660  
DB |||||  
601 GAAAAAATTTCAAAAATTTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 660  
QY 661 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAATTAACACTACT 720  
DB |||||  
661 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAATTAACACTACT 720  
QY 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTT 780  
DB |||||  
721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTT 780  
QY 781 AAAAAATTAATAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
DB |||||  
781 AAAAAATTAATAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTTTCAAGGAAAAATAA 876  
DB |||||  
841 CCTAGTGATATAAGCATTTTCAAGGAAAAATAA 876

RESULT 13

AAAS3725

ID AAAS3725 standard; DNA; 876 BP.

XX AC AAAS3725;

XX AC AAAS3725;

DT 15-SEP-2003 (revised)

XX DT 22-DEC-2000 (first entry)

DE DT 22-DEC-2000 (first entry)

XX Campylobacter jejuni O:41 serotype alpha-2,3-sialyltransferase.

KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside; ds.  
XX Campylobacter jejuni; O:41 serotype.  
OS  
FH Key Location/Qualifiers  
FT 1..876  
FT /\*tag= a  
FT /product= "alpha-2,3-sialyltransferase"  
XX  
PN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CA000086.  
XX  
PR 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
PA (CANA ) NAT RES COUNCIL CANADA.  
XX  
PI Gilbert M, Wakarchuk WW;  
XX  
DR WPI; 2000-524418/47.  
DR P-PSDB; AAY97211.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
PS Claim 6; Page 94; 120pp; English..  
XX

CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Query Match 93.6%; Score 820; DB 3; Length 876;

Best Local Similarity 96.0%; Pred. No. 5.9e-125;

Matches 841; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 ATCAAAAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60

DB 1 ATGAAAAAGTTATTATTGCTCGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 60

QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120

DB 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTGTAAGATAAATACTAT 120





Search completed: July 19, 2006, 15:26:26  
Job time : 657.496 secs

---

GenCore version 5.1.9  
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:59:51 ; Search time 4168.52 Seconds  
(without alignments)  
11751.261 Million cell updates/sec

Title: US-10-734-719-8

Perfect score: 876

Sequence: 1 atgaaaaaagtattattgc.....attatttcaaggaaaaataa 876

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est3:\*  
3: gb\_est4:\*  
4: gb\_est5:\*  
5: gb\_est6:\*  
6: gb\_est7:\*  
7: gb\_est2:\*  
8: gb\_est7:\*  
9: gb\_est8:\*  
10: gb\_est9:\*  
11: gb\_gss1:\*  
12: gb\_gss2:\*  
13: gb\_gss3:\*  
14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	10.6	1101	14	CNS00EVL
2	92	10.5	1359	9	DN685273
3	92	10.5	1365	10	DT959930
4	90.8	10.4	1179	1	AJ927794
5	90.4	10.3	1007	9	DN566957
6	89.4	10.2	1101	14	CNS0039G
7	88.4	10.1	1227	14	AG430010
8	87.8	10.0	1542	14	AG386981
9	86.4	9.9	1210	12	CG749728
10	85.8	9.8	1324	14	AG376784
11	85.8	9.8	1354	12	CG744717
12	85.6	9.8	976	13	CW949642
13	85.6	9.8	1096	1	AJ927821
14	85.4	9.7	1297	10	DV778934
15	84.4	9.6	947	12	CL112501
16	84.4	9.6	1101	14	CNS00801
17	84.2	9.6	1348	12	CG749499
18	84	9.6	1364	12	CG757970
19	83.6	9.5	734	14	CNS010MP

C	20	83.4	9.5	1123	5	CD325190
C	21	83.2	9.5	906	14	AG555384
C	22	83.2	9.5	1544	10	DV780516
C	23	83	9.5	1225	14	CNS0161D
C	24	83	9.5	1242	12	CL068807
C	25	82.4	9.4	930	7	BB898645
C	26	82.4	9.4	1248	9	DN808090
C	27	82.2	9.4	963	10	DT797234
C	28	82.2	9.4	1101	14	CNS00EVL
C	29	81.8	9.3	759	14	CNS06QXV
C	30	81.8	9.3	1133	1	AJ926090
C	31	81.8	9.3	1638	14	AG347262
C	32	81.4	9.3	1092	1	AJ927993
C	33	81.4	9.3	1258	13	CL508886
C	34	81.4	9.3	1268	14	AG347098
C	35	81.4	9.3	1489	10	DT957227
C	36	81.2	9.3	1190	1	AJ927689
C	37	81	9.2	1277	12	CC253231
C	38	80.8	9.2	1592	12	CG750135
C	39	80.6	9.2	1007	9	DN566957
C	40	80.6	9.2	1118	10	DM600040
C	41	80.6	9.2	1140	12	CL071552
C	42	80.6	9.2	1186	13	CL649642
C	43	80.4	9.2	1227	10	DT989558
C	44	80.4	9.2	1355	14	AG346348
C	45	80	9.1	949	5	CK424488

## ALIGNMENTS

### RESULT 1

#### CNS00EVL

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### COMMENT

#### FEATURES

#### source

1. 1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR29B23"  
/clone\_lib="RPCI-98"

CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999  
Drosophila melanogaster genome survey sequence T7 end of BAC:  
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL069706

AL069706.1 GI:4949849

GSS:

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammeter in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

pi and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### Location/Qualifiers

1. 1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR29B23"  
/clone\_lib="RPCI-98"







```

Db      572 ATAATATATATATTTTATAATATATATTTATTTATATAATAATAATAATAATAATAATAATA 513
Qy      807 TTTATACCAAGTTGATAAAAAGATCTATTAAAGATTACCTAGTGATATAATAAGCATTATTTTCAA 866
Db      512 TTAATTTTATATTTTAAATTTATTTTAAATATATATATATATATATATATTTTAAATAATAATAA 453
Qy      867 A 867
Db      452 A 452

RESULT 4
AJ927794
LOCUS
DEFINITION
AJ927794 Theileria annulata merozoite Theileria annulata cDNA clone
tam03h06_plk, mRNA sequence.
ACCESSION
AJ927794
KEYWORDS
AJ927794.1 GI:67498179
EST.
Theileria annulata
Theileria annulata
Theileria annulata
Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
Theileria.
1 (bases 1 to 1179)
REFERENCE
AUTHORS
Pain, A., Renauld, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W.,
Kerhornou, A., Aslett, M., Bishop, R., Bouchier, C., Cochet, M.,
Coulson, R.M.R., Cronin, A., de Villiers, E., Fraser, A., Foster, N.,
Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F.,
Larke, N., Lord, A., Maser, P., McKellar, S., Mooney, P., Morton, F.,
Nene, V., O'Neill, S., Price, C., Quail, M.A., Rabinowitsch, E.,
Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T.,
Squares, R., Squares, S., Tivey, A., Walker, A.R., Woodward, J.,
Dobbelare, D.A.E., Langsley, G., Rajandream, M.-A., McKeever, D.,
Shiels, B., Tait, A., Barrell, B. and Hall, N.
The genome of the host-cell transforming parasite Theileria
annulata and a comparison with T. parva
Unpublished (2005)
JOURNAL
COMMENT
Contact: Pain A
The Pathogen Sequencing Unit
The Wellcome Trust Sanger Institute
Genome Campus, H10 1SA, UNITED KINGDOM
Merozoite cDNA library: Frank Katzer and Brian Shiels, Division of
Veterinary Infection and Immunity, ICM, University of Glasgow, UK.
FEATURES
source
1. 1179
/organism="Theileria annulata"
/mol_type="mRNA"
/isolate="Ankara (clone D7)"
/db_xref="taxon:5874"
/clone="tam03h06_plk"
/dev_stage="merozoite"
/lab_host="Bos taurus (cow)"
/clone_libs="Theileria annulata merozoite"
/note="country: Turkey; Ankara"

ORIGIN
Query Match 10.4%; Score 90.8; DB 1; Length 1179;
Best Local Similarity 44.0%; Pred. No. 2.2e-05;
Matches 281; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 238 AATTACACCAAGCTCATCTAGAAATCGAAATTTTGTAAAAACCTTTTACGATTATTTT 297
Db 332 ANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 391
Qy 298 CCTGATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTTTAATGCTTATTTT 357
Db 392 NATAAAAATTTATNTATTTTAAATATATTTTNNATATAATAAAAAAAATATTATTAATAAT 451
Qy 358 AAATTTACGAAATTTATTTCAATCAAGAATTCCTCAGGGCTCTATATGTGTCAGTA 417
Db 452 AAAAAAAATTAATAATTTTAAATATATATTAATTAATAATATTTTAAAAATAATAATA 511

```

```

Qy 418 GCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGG 477
Db 512 AAAATAATAAAATTTATAAAATATTAATAATAATAATAATAATAATAATAATAATAAT 571
Qy 478 TCATCTTATGCTTTTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAA 537
Db 572 TAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 631
Qy 538 AATGATCGCTCGACATATATCGACATAGTAAATAATACAGATATATAAAGCTTTTGAATTT 597
Db 632 ATATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 691
Qy 598 CTAGAAAAAATTTACAAAATAAACTATATTCGTTATGTCCTAATAGCTTTTTCAGCAAT 657
Db 692 ATATATTTAAAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 751
Qy 658 TTTATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTAC 717
Db 752 TTTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 811
Qy 718 ACTAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAAT 777
Db 812 AAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 871
Qy 778 TTTAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 837
Db 872 TAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 931
Qy 838 TTACCTAGTGATATATAAGCAATTTTCAAGGAAAAATAA 876
Db 932 TTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 970

RESULT 5
DN566957
LOCUS
DEFINITION
92914460 Sea Urchin primary mesenchyme cell cDNA library
Strongylocentrotus purpuratus cDNA clone PMCSPR2-143H12 3', mRNA
sequence.
ACCESSION
DN566957
VERSION
DN566957.1 GI:61125996
KEYWORDS
EST.
SOURCE
Strongylocentrotus purpuratus
ORGANISM
Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoidea;
Strongylocentrotidae; Strongylocentrotus.
REFERENCE
1 (bases 1 to 1007)
AUTHORS
Zhu, X., Mahairas, G., Illies, M.R., Cameron, R.A., Davidson, E.H. and
Ettensohn, C.A.
TITLE
A large scale analysis of mRNAs expressed by primary mesenchyme
cells of the sea urchin embryo
JOURNAL
PUBMED
Development 128 (13), 2615-2627 (2001)
11493577
COMMENT
Contact: Erica Sodergren
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-7676
Fax: 713-798-6977
Email: ericas@bcm.tmc.edu
NCBI Trace Archive: 496925179
Insert Length: 1750 Std Error: 0.25
Plate: 143 row: H column: 12.
Location/Qualifiers
1. 1007
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="PMCSPR2-143H12"
/tissue_type="embryo"
/cell_type="primary mesenchyme cells"
/lab_host="E.coli"

```

		/clone_lib="Sea Urchin primary mesenchyme cell cDNA library" /note="Vector: pSPORn1; Site 1: NotI; Site 2: MluI/ oligo dt priming from poly A+ RNA_directionally_cloned"	
ORIGIN			
	Query Match Best Local Similarity Matches 394; Conservative	10.3%; Score 90.4; DB 9; Length 1007; 46.9%; Pred. No. 2.6e-05; 0; Mismatches 431; Indels 15; Gaps	
QY	36	TTTAAAGAAATGATTATTCAGGCTACCAACTGATTTTGATGTATTAGATGTAATCA	95
DB	165	TTTTAAAAATTTTTTTTAAAAAATAAAAAAAAAAATTTAAAAAATAATTTAA	224
QY	96	ATTTTATTTGAAGATAAATACTATCTGGTAAAAAATGCAAAGCAGTGTTTTACACCC	155
DB	225	AAAATTTTTTTTAAAAAATAAAAAAAAAAATTTTTTTTTTTTAAAAAATA	284
QY	156	TAAATTTCTTTGAGCAATACACACTTTAAAAACATTTAATCCAAATCAAGAATATCA	215
DB	285	TTTTTTTTTTTTTTAAAAAAGGGTTTTTTTTTTTTTGTGTAATAATTTTAAAAA	344
QY	216	GACCGAACATTAATATGTCCTTAATACAACCAAGCTCATCTAGAAAAATGAAATTTGT	275
DB	345	AAACCTTTTTTTTTTTTTTTTATAAAAAATAAAAAAAAAATTTTTTTTTTAAAAA	404
QY	276	AAAAACTTTTTACGATTAATTTCTGATGCTCATTTGGGATATGATTTTTTAAACAACT	335
DB	405	AAAAATTTTATTTAAAAAATTTAAAAATTTTTTTTTTAAATAAATTTTTTTAAAAAT	464
QY	336	TAAAGAAATTAATGCTTAATTTTAAATTTTCAGAAATTTAATTCAAAGAAATPACCTC	395
DB	465	CCAAAAAATAATAATTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTAAAAATTTT	524
QY	396	AGGGGTCTATATGTCGAGTAGCCATAGCCTAGGACATACAAAGAAATTTATCTTCGG	455
DB	525	AA-----TTTAAAAATTTATTTTATTTATTTATATAAAAAAATAATTTTAAAAATA	579
QY	456	AATGTATTTTATCAAATGGTGCATCTTATGCTTTTGATACCACCAAGAAATCTTTT	515
DB	580	TTTAAATTTTAAAAAATAAATAATTTTTTTTTTTTTTTTTTTTTTTTAAAAATTTTTT	639
QY	516	AAAACTAGCCCCTGATTTTAAAAATGATPCGCTCGCACATATATCGGACATAGTAAAAATAC	575
DB	640	TTTTTTATTTTTTCTTCTAAAAAATA-----AAAAATTTTAAATTTAGAAAAATTA	691
QY	576	AGATATAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAATAAATCTATATGCTTTATG	635
DB	692	AAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	749
QY	636	TCCTAAATAGTCTTTTAGCAATTTTATAGAACTAGCGCCAATTTAAATTCAAATTTTAT	695
DB	750	TTTTTTTTTTTTTTTAAAAATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA	809
QY	696	CATACAGAAAAAATAACTACACTAAAGATATATCTCATCTCTAGTGAGGCTTAATGG	755
DB	810	ATATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	869
QY	756	AAAATTTTCAAAAAATTAATTTTAAAAAATAAATAAATAAAGAAATGTTTATATCAAA	815
DB	870	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	929
QY	816	GTTGATAAAGATCTTATTAAGATTACCTAGTGATATAAGCATTTATTTCAAGGAATA	875
DB	930	TTTCCTTTTATTTTATTTTATTTTTTTTTTAAAAAATAAATAAATAAATAAATAAATAA	989
RESULT 6			
CNS0039G		1101 bp DNA linear	GSS 03-JUN-1999
LOCUS		Drosophila melanogaster genome survey sequence TET3 end of BAC #	
DEFINITION		BACR0BK10 of RPCL1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

**COMMENT**

FEATURES  
source

## ORIGIN

Query Match 10.2%; Score 89.4; DB 14; Length .1101;  
Best Local Similarity 18.8%; Pred. No. 3.9e-05;  
Matches 132; Conservative 299; Mismatches 270; Indels 2; Gaps 1;

QY	70	GA	TTTTGATG	TATTTAG	TGTAATCA	AAATTTAT	TTTGA	AGATAAA	TACTATCT	TGGTAAA	129
Db	394	GM	ATAWAAW	WWWWTTTT	TTTAA	WAAAAATA	TAATTTW	AAWAAAAA	AAAAATTT	WAAAAWA	453
QY	130	AA	TGCAAG	CAGTGT	TTTAC	ACCCCTAA	TTTCTT	CTTGAG	CAATACT	ACACTTTAAA	189
Db	454	AA	AWATW	AWTTT	TAWT	WAAAAA	AAAAAAT	TTTTTTTT	TTTAA	TTTAAWTTT	513
QY	190	CA	TTTAAAT	CCAAAT	CAAGA	TATGAG	CCGAAC	TAAAT	TATGTTCT	TAAATCAACCA	249
Db	514	AA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	TAAAT	TWTWTT	TTTWWAAAT	573
QY	250	GC	TATCTAG	AAAAAT	TTTGT	AAAAAC	TTTTT	TACGA	TATTTCT	TCTGATGCTCAT	309
Db	574	TY	TTYHY	TYTY	TWTW	HYHTMY	THA	WHTTW	HYHTY	AWMHWTW	633
QY	310	TT	GGAATAT	GATTT	TTTAA	CAAC	CTTAA	AGAAAT	TTTAA	TGCTTATTTAAAT	369
Db	634	TY	TAAYYY	YTY	CMYY	HYHMH	HAHA	AAAWTT	TTTW	THAVHWATY	693
QY	370	AT	TATTTCA	ATCA	AGAAT	TACCT	CAGGGT	CTATAT	GTGTG	CAGTAGCCATAG	429
Db	694	CT	HTCHC	YHHY	HTAH	HTHTH	HHY	HAHY	MYMY	WYMYTCTACT	751
QY	430	GG	ATACA	AGAAAT	TTATCT	TCGGGA	TTGAT	TTTTT	TATCA	AAATGGGT	489
Db	752	A	HTTW	YAWAH	AMW	MHH	HAHY	AAAAA	AAWAT	HTHHY	811
QY	490	TT	TGATC	CAAAA	CAAGAA	ATCTTT	TTTAA	ATCAG			

RESULT 6	CNS0039G	DNA	linear	GSS 03-JUN-1999
CNS0039G		1101 bp		
LOCUS	CNS0039G			
DEFINITION	Drosophila melanogaster: genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			



Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and Shiroishi,T.  
Contribution of Asian mouse subspecies *Mus musculus molossinus* to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis  
Genome Res. 14 (12), 2439-2447 (2004)  
15574823  
2 (bases 1 to 1542)  
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.  
Direct Submission  
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan  
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan  
(E-mail:hattori@gsc.riken.jp, URL:http://hgsc.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).  
Tsukuba Institute, Bio Resource Center,  
The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@rtc.riken.jp  
PRIMERS  
Sequencing : TJ  
LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.  
Location/Qualifiers  
1..1542  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSMg01-201G10.TJ"  
/sex="male"  
/tissue\_type="mixture of kidney and spleen"  
/clone\_lib="MSMg01 Mouse Male BAC Library"

Query Match 10.0%; Score 87.8; DB 14; Length 1542;  
Best Local Similarity 46.8%; Pred. No. 6.9e-05;  
Matches 411; Conservative 0; Mismatches 453; Indels 14; Gaps 5;

QY 5 AAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAGGCTAC 64  
DB 1441 ATNAAATAAANTANTANTATATATATTAAGATATATAAATATATATATATATAN 1382

QY 65 CAATGATTGATGATTAGATGTAATCAATTTTATTTTGAAGATAAATCTATCTTG 124  
DB 1381 AAAATATTAAATAATAAANNATAGTAAATAATTAATAAATAATATATATAATNA 1322

QY 125 GTAAAAATGCAACAGTGTTTACACCCCTAAATTCCTTTGAGCAATCTACAGTT 184  
DB 1321 ATTAATAATATATA-TATAATAATAAAATATATAATTTTANAATAATAAAATAA 1263

QY 185 TAAACATTTTAAATCCAAATCAAGATATGAGACCGAATATTTATGCTTCTAATTACA 244  
DB 1262 TAATAATAAATAATATAATATAATATTAATAATAAANAATAAATAAATAATNA 1206

QY 245 ACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTTCCTGATG 304  
DB 1205 AAATAATAATATATATATATATATATATTAATAAATAAATAAATAAATAAATAA 1146

QY 305 CTCATTGGGATATGATTTTAAACAACTTAAGAAATTTAATGCTTATTTTAAATTC 364  
DB 1145 TTTAAATAAATAATATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATA 1086

QY 365 ACGAAATTTTATTCATCAAGAAATTCCTCAGGGGTCTATATG-TGTGCACTAGCCATA 423  
DB 1085 AAATAATAAATAAATAATATTAATAAATAAATAAATAAATAAATAAATAAATAA 1026

QY 424 GCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTGATTTTATCAAAATGGGTCACT 483

Db 1025 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 970  
QY 484 TATGCTTTTGCATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAATGAT 543  
DB 969 TATAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 910  
QY 544 CGCTCGCATTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTGAATTTCTAGAA 603  
DB 909 TATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 850  
QY 604 AAAACCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 663  
DB 849 AAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 790  
QY 664 GAACCTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAA-----AAAAATAACTACA 718  
DB 789 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 730  
QY 719 CTAAGATATATCTCATACCTTCTTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATT 778  
DB 729 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 670  
QY 779 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 838  
DB 669 TTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 610  
QY 839 TACCTAGTGATATAAAGCATTTTCAAGAGAAATAA 876  
DB 609 AATATTAAATAATTTAATAATNTAATTTAATAAATAA 572

RESULT 9  
CG749728/c

LOCUS  
DEFINITION

CG749728  
ACCESSION

CG749728.1 GI:37970654  
VERSION

GSS.  
KEYWORDS

Pristionchus pacificus  
SOURCE

Pristionchus pacificus  
ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.

REFERENCE  
AUTHORS

1 (bases 1 to 1210)  
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
Buntjer,J., van der Meulen,M. and Sommer,R.J.

TITLE  
JOURNAL

An integrated physical and genetic map of the nematode *Pristionchus pacificus*  
Mol. Genet. Genomics 269 (5), 715-722 (2003)

PUBMED  
COMMENT

Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de

Class: BAC ends.  
Location/Qualifiers

1..1210  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"

/db\_xref="taxon:54126"  
/clone\_lib="Ppa EcoRI BAC Library"

/notes="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

ORIGIN

Query Match 9.9%; Score 86.4; DB 12; Length 1210;  
Best Local Similarity 40.8%; Pred. No. 0.00012;

[illegible]

RESULT 10	AG376784/c	AG376784	1324 bp	DNA linear	GSS 21-DEC-2004
LOCUS		Mus musculus molossinus	DNA, clone:MSMg01-184T09.T7,		genomic survey
DEFINITION		sequence.			
ACCESSION		AG376784			
VERSION		AG376784.1	GI:47987989		
KEYWORDS		GSS.			
SOURCE		Mus musculus molossinus	(Japanese wild mouse)		
ORGANISM		Mus musculus molossinus			

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
AUTHORS	1 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and Shirosishi, T.
TITLE	Contribution of Asian mouse subspecies <i>Mus musculus molossinus</i> to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis
JOURNAL	Genome Res. 14 (12), 2439-2447 (2004)
PUBMED	15574823
REFERENCE	2 (bases 1 to 1324)
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
JOURNAL	1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
	Clones are derived from the mouse BAC library M5Mg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp), Tsukuba Institute Bio Resource Center,
COMMENT	The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp

**PRIMERS**  
 Sequencing : T7  
**LIBRARY**  
 Vector : pBACe3.6  
 R.Site 1 : EcoRI  
 R.Site 2 : EcoRI.

FEATURES	source
k.site 2	: ECCR1
Location/Qualifiers	
1..1324	
/organism="Mus musculus molossinus"	
/mol_type="genomic DNA"	
/sub_species="molossinus"	
/db_xref="taxon:57486"	
/clone="MSMG01-184109.T7"	
/sex="male"	
/tissue_type="mixture of kidney and spleen"	
/clone_lib="MSMG01 Mouse Male BAC Library"	

<b>ORIGIN</b>					
		Query Match	9.8%;	Score 85.8;	DB 14; Length 1324;
		Best Local Similarity	46.0%;	Pred. No. 0.00015;	
		Matches 324:	Conservative	0;	Mismatches 375; Indels 5; Gaps 1;

[illegible]

QY 533 TTAATAATGATCGCTCGCACTATATCGGCATATAGTAAATAACAGATATAAAAGCTTTAG 592  
 |||||  
 Db 912 AAAAAATAATAATATAAAAAATTT-----AAATTAATNATAAATAATAAATAAAAA 858  
 QY 593 AATTTCTAGAAAAAATTTCAAAATAAATACTATATTTGCTTTATGTCCTTAATAGTCTTTTAG 652  
 |||||  
 Db 857 AATTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 798  
 QY 653 CAATTTTATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAAATA 712  
 |||||  
 Db 797 ATAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 738  
 QY 713 ACTACACTAAGATATACCTATACCTTCTAGTCAGGCTTATGCGAAATTTTCAAAATA 772  
 |||||  
 Db 737 TATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 678  
 QY 773 TTAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 832  
 |||||  
 Db 677 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 618  
 QY 833 TAAGATTACCTAGTATATAAGCATATTTTCAAGGAAATAA 876  
 |||||  
 Db 617 ATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 574

RESULT 11  
 CG744717/c  
 LOCUS  
 DEFINITION  
 P037-2-P03.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,  
 genomic survey sequence.  
 ACCESSION  
 CG744717  
 VERSION  
 CG744717.1 GI:37965585  
 SOURCE  
 Pristionchus pacificus  
 ORGANISM  
 Pristionchus pacificus  
 Pristionchus pacificus  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
 Neodiplogasteridae; Pristionchus.  
 REFERENCE  
 1 (bases 1 to 1354)  
 Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
 Buntjer,J., van der Meulen,M. and Sommer,R.J.  
 An integrated physical and genetic map of the nematode Pristionchus  
 pacificus  
 JOURNAL  
 Mol. Genet. Genomics 269 (5), 715-722 (2003)  
 PUBLISHED  
 12884007  
 COMMENT  
 Contact: Sommer RJ  
 Evolutionary Biology  
 Max-Planck-Institute for Developmental Biology  
 Spemannstr. 37-39, Tuebingen D-72076, Germany  
 Tel: 00497071601371  
 Fax: 00497071601498  
 Email: ralf.sommer@tuebingen.mpg.de  
 Class: BAC ends.

FEATURES  
 Location/Qualifiers  
 1..1354  
 /organism="Pristionchus pacificus"  
 /mol\_type="genomic DNA"  
 /strain="California"  
 /db\_xref="taxon:54126"  
 /clone\_lib="Ppa EcoRI BAC Library"  
 /note="The library was generated by a partial digest of  
 the genomic DNA with EcoRI and cloning into the BAC  
 vector."  
 ORIGIN  
 Query Match 9.8%; Score 85.8; DB 12; Length 1354;  
 Best Local Similarity 46.7%; Pred. No. 0.00015;  
 Matches 390; Conservative 0; Mismatches 437; Indels 8; Gaps 4;  
 QY 4 AAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAGAAATTCATTATTCAAGGCTA 63  
 |||||  
 Db 1319 AAAATAAATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1260  
 QY 64 CAAATGATTTTGATGTTAGATGTAATCAATTTTATTTTGAAGATAAATACTATCTT 123

Db 1259 AAAAAATAAT 1200  
 |||||  
 QY 124 GGTAAAAAATGCAAGCAGTGTCTTTACACCCCTAAATTTCTCTTTGAGCAACTACTACACT 183  
 |||||  
 Db 1199 AAATAAATAAT 1140  
 |||||  
 QY 184 TTAACAATTTAATCCAAAAATCAAGAAATATGAGACCGAACTAATATATGTTGTTCTTAATTAC 243  
 |||||  
 Db 1139 AAATTTATATAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1084  
 |||||  
 QY 244 AACCAAGCTCATCTAGAAATGCAAAATTTTGTAAACCTTTTACGATATTTTCTCTGAT 303  
 |||||  
 Db 1083 AAATAAATTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1024  
 |||||  
 QY 304 GCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAAATTT 363  
 |||||  
 Db 1023 ATAAAT 965  
 |||||  
 QY 364 CAGAAATTTTAAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCCATA 423  
 |||||  
 Db 964 TATTATAAATATTTATAAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 905  
 |||||  
 QY 424 GCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATCGGTCATCT 483  
 |||||  
 Db 904 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 845  
 |||||  
 QY 484 TATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAATGAT 543  
 |||||  
 Db 844 AAAAAAT 785  
 |||||  
 QY 544 CGCTCGCATATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAA 603  
 |||||  
 Db 784 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 725  
 |||||  
 QY 604 AAAACTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 663  
 |||||  
 Db 724 AAATTTATTCGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 665  
 |||||  
 QY 664 GAATCGCGCAAAATTTAAAT-TCAAAATTTTATCATCAAGAAATAAATAAATAAATAAATAAATA 722  
 |||||  
 Db 664 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 605  
 |||||  
 QY 723 AGATATCTCATATACCTTCTAGTAGGCTTATGAAAAAT--TTTCAAAAAATATTAATTTT 780  
 |||||  
 Db 604 ATATTTATTTTATTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 545  
 |||||  
 QY 781 AAAAAATAAATAAATAAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAA 835  
 |||||  
 Db 544 ANTAATTTTATAAATTTTAAATAATTTATTAATTAATTTTATTAATTTATTTATTTAA 490  
 |||||

RESULT 12  
 CW949642  
 LOCUS  
 DEFINITION  
 TcB34.3 D04 SP6 Tribolium BAC library Tribolium castaneum genomic,  
 genomic survey sequence.  
 ACCESSION  
 CW949642  
 VERSION  
 CW949642.1 GI:56732262  
 SOURCE  
 Tribolium castaneum (red flour beetle)  
 ORGANISM  
 Tribolium castaneum  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
 Tenebrionidae; Tribolium.  
 REFERENCE  
 1 (bases 1 to 976)  
 Savard,J. and Tautz,D.  
 Tribolium castaneum BAC-ends sequencing project  
 UNPUBLISHED (2003)  
 CONTACT: Savard, J.  
 Abteilung für Evolutionsgenetik, AG Tautz  
 Institut für Genetik, Universität zu Köln  
 Weyertal 121, 50931 Köln, Germany



```

Db      607 TAAATATATAAATANNAAAAATNTAANTAANNAANNAANNTTANAATTTAATAA 666
QY      520 CTAGCCCTGATTTTAAATAATGATCGCTCGCACTATATCGACATAGTAAAAATACAGAT 579
Db      667 TTANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 726
QY      580 ATAAAGCTTTAGATTTCTAGAAAACCTTCAAAAATAAACTATATATGCTTATGTCCT 639
Db      727 NAANANNNNNNANNAT-----AAAANNNNAATTAATAAATAAATAAATAAATAAATAA 780
QY      640 AATAGTCTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATA 699
Db      781 AANNAATTTANAANNNTTTAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 840
QY      700 CAAGAAAAAATAAATACATACTAAAGATATATCTCATACCTTCTAGTGAGGCTTATGAAAA 759
Db      841 AAAAAAANTANNATNTNNNAATTANNATNAAATNNNAANNAANNAANNAATAGTATCTANTAA 900
QY      760 TTTTCAAAAAATATTAATTTTAAAAAATAAATAAATAAAGAAATGTTTATTAACAAGTTG 819
Db      901 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
QY      820 ATAAAGATCTATTAAGATTTACCTAGTGATATAAGCATTTATTCAAAGGAAATAA 876
Db      961 ANNAANATATANNAAAAAATAATTAATAATTAATAAATAAATAAATAAATAAATAAATAA 1017

RESULT 14
LOCUS   DV778934
DEFINITION Hw Fat_9 C11 Bos taurus CF-24-HW fat cDNA library Bos taurus CDNA, mRNA sequence. EST 23-NOV-2005
ACCESSION DV778934.1 GI:82631810
VERSION   DV778934
KEYWORDS  EST.
SOURCE    Bos taurus (cattle)
ORGANISM  Bos taurus
REFERENCE Yoon,D.H., Lee,S.H., Park,S.H., Park,E.W., Cho,Y.M., Lee,J.H., Kim,H., Kim,H.Y., Park,J.H. and Oh,S.J.
AUTHORS   Gene Expression Profiling of the Bovine liver, adipose, and skeletal muscle
TITLE     Unpublished (2005)
JOURNAL   Contact: Dr. Du-Hak Yoon
COMMENT   National Livestock Research Institute, RDA
          564 Omoekchun-dong, Suwon, 441-350, Korea
          Tel: 82 31 290 1593
          Fax: 82 31 290 1792
          Email: dhyoon@rda.go.kr.

FEATURES             Location/Qualifiers
     source           1..1297
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /sex="Four males mixed"
                     /cell_type="adipocyte"
                     /dev_stage="24 months old"
                     /lab_host="XLI-BlueMRF" strain"
                     /clone_lib="Bos taurus CF-24-HW fat cDNA library"
                     /note="Organ: adipose tissue; Vector: Uni-ZAPXR; Site_1: ECORI; Site_2: Xho I"

ORIGIN
Query Match          9.7%; Score 85.4; DB 10; Length 1297;
Best Local Similarity 47.3%; Pred. No. 0.00018;
Matches 330; Conservative 0; Mismatches 361; Indels 7; Gaps 3;

QY      186 AAAACATTTATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTTAATACAA 245
          ||||| + || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      466 AAAAAAAAAAANATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTCAT 525
QY      246 CCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTAATTTTCCCTGATGC 305
Db      526 ATTTTATTTTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 585
QY      306 TCATTTGGGATATGATTTT---TTTAAACAACCTTAAAGAAATTTAATGCTTATTTAAATTT 362
Db      586 AATTTAATAAATAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 645
QY      363 TCAGGAATTTATTTCAATCAAGAAATTTACCTCAGGGCTCTATATGTGCGAGTAGCCAT 422
Db      646 AATTNAATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 705
QY      423 AGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTATCAAAATGGGTCATC 482
Db      706 AAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 765
QY      483 TTATGCTTTTGTATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAATGA 542
Db      766 AAAAAAATTTTAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATTT 825
QY      543 TCCTCGCACTATATCGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTTACA 602
Db      826 TAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 885
QY      603 AAAAATCTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 660
Db      886 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 945
QY      661 ATAGAACTAGCGCAAAATTTTAAATTTCAAAATTTTATCATACAGAAAAAATAAATAAATA 720
Db      946 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1005
QY      721 AAAGATATATCTATACCTCTCTAGTGAGCTTATGGA--AAATTTCAAAAAATAATAATTT 778
Db      1006 TTTTAAAAAATTTATNTTTTATATAATTTTAAATNANNNTTTTAAAAAATAAATAAATA 1065
QY      779 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 838
Db      1066 TTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1125
QY      839 TACTAGTGATATAAAGCATTTTTCAAAGGAAAAATAA 876
Db      1126 TAATAATTTTATATTTTATTTTAAAAATAAATAA 1163

RESULT 15
LOCUS   CL112501/c
DEFINITION ISB1-57D23_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-57D23, genomic survey sequence.
ACCESSION CL112501
VERSION   CL112501.1 GI:40606136
KEYWORDS  GSS.
SOURCE    Xenopus tropicalis (western clawed frog)
ORGANISM  Xenopus tropicalis
REFERENCE 1 (bases 1 to 947)
AUTHORS   Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE     A physical map of the xenopus tropicalis genome
JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K Wilson
          Genome Sequencing Center
          Washington University School of Medicine
          Email: submissions@watson.wustl.edu
          Insert length: 75000 Std Error: 0.00
          Seq primer: T7 TAATACGACTCACTATAGG
          Class: BAC ends
          High quality sequence start: 3

```



```
FEATURES             High quality sequence stop: 334.
Location/Qualifiers
  1..947
    /organism="Xenopus tropicalis"
    /mol_type="genomic DNA"
    /db_xref="taxon:8364"
    /clone="ISB1-57D23"
    /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis; BAC
Library Segment 1"

ORIGIN
Query Match      9.6%; Score 84.4; DB 12; Length 947;
Best Local Similarity 45.8%; Pred. No. 0.00027;
Matches 260; Conservative 0; Mismatches 302; Indels 6; Gaps 1;

Qy 309 TTTCGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAATTTTCACGA 368
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 946 TTTTATAAATTTTATATTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 887
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 369 AATTATTTCATCAAGAATTAACCTCAGGGGTCTATATGTGTCAGTAGCCATAGCCCT 428
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 886 AATTTTTTTTATAATAAATATTTTAAATAATTTTATAATAATTTTATAATAATTTTAAAT 827
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 429 AGGATACAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGTCACTTATGC 488
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 826 ATAAAAAATAAAATTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 767
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 489 TTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAATGATCGCTC 548
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 766 ATTTTTTTATAAAAAATATTTTAAAAAAATTTTATAAAAAATTTTATAAAAAATTTTAA 707
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 549 GCATATATCGGCATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTAGAAAAAC 608
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 706 TTATATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 647
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 609 TTACAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAATTTTATAGAACT 668
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 646 TTAATTTAAAAATAAAATTTTATTATTNAAAATTTATATTAAATAAATAAATAAATAAATA 590
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 669 AGCCCAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAAATAAATAAATAAATA 728
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 589 ---NNAAAAATAAANAATAATTTTAAANAATAAANAATAAANAATAAANAATAAANAATA 533
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 729 ACTCATACCTCTCTAGTCAGGCTTATGGAATTTTCAAAAAATATTAATTTTAAAAAAT 788
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 532 AANNAAAAAATAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATA 473
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 789 AAAAAATTAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTACCTAGTGA 848
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 472 AATTTAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATA 413
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 849 TATAAGCATTTTTCAAAGGAATAA 876
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 412 AAAAAAATAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATAAANAATA 385
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 19, 2006, 22:40:45  
Job time : 4171.52 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:55:46 ; Search time 5067.9 Seconds  
(without alignments)  
11053.473 Million cell updates/sec

Title: US-10-734-719-8  
Perfect score: 876  
Sequence: 1 atgaaaaagttatttgc.....attatttcaaggaaataa 876

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_env.\*
- 2: gb\_pat.\*
- 3: gb\_ph.\*
- 4: gb\_pl.\*
- 5: gb\_pr.\*
- 6: gb\_ro.\*
- 7: gb\_sts.\*
- 8: gb\_sy.\*
- 9: gb\_un.\*
- 10: gb\_vi.\*
- 11: gb\_ov.\*
- 12: gb\_hcg.\*
- 13: gb\_in.\*
- 14: gb\_on.\*
- 15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	876	100.0	876	2. AX934431	Sequence
2	876	100.0	876	2. BD249794	Campyloba
3	876	100.0	876	2. AR271703	Sequence
4	876	100.0	876	2. AR481785	Sequence
5	876	100.0	876	2. AR527384	Sequence
6	876	100.0	876	2. AR609663	Sequence
7	876	100.0	876	2. AR689941	Sequence
8	876	100.0	876	2. AR691837	Sequence
9	876	100.0	6047	15. AY661458	Campyloba
10	876	100.0	24425	15. AY422197	Campyloba
11	876	100.0	24437	15. AF167344	Campyloba
12	874.4	99.8	4749	15. AF305571	Campyloba
13	872.8	99.6	12390	15. AF401528	Campyloba
14	869.8	99.3	873	2. AX934436	Sequence
15	869.8	99.3	873	2. AR481787	Sequence
16	869.8	99.3	873	2. AR527386	Sequence
17	869.8	99.3	873	2. AR609665	Sequence
18	869.8	99.3	873	2. AR689943	Sequence

19	859.8	99.3	873	2	AR691839	Sequence
20	859.6	99.3	11474	15	AF215659	Campyloba
21	856.6	98.9	873	2	AX934434	Sequence
22	856.6	98.9	873	2	AR481786	Sequence
23	856.6	98.9	873	2	AR527385	Sequence
24	856.6	98.9	873	2	AR609664	Sequence
25	856.6	98.9	873	2	AR689942	Sequence
26	856.6	98.9	873	2	AR691838	Sequence
27	852.4	97.3	1303	15	AY681242	Campyloba
28	844	96.3	876	2	AX934427	Sequence
29	844	96.3	876	2	BD249792	Campyloba
30	844	96.3	876	2	AR271701	Sequence
31	844	96.3	876	2	AR481783	Sequence
32	844	96.3	876	2	AR527382	Sequence
33	844	96.3	876	2	AR609661	Sequence
34	844	96.3	876	2	AR689939	Sequence
35	844	96.3	876	2	AR691835	Sequence
36	842.4	96.2	11442	15	AF400048	Campyloba
37	837.6	95.6	876	15	AF216647	Campyloba
38	837.6	95.6	6047	15	AY644679	Campyloba
39	837.6	95.6	11474	2	AX934424	Sequence
40	837.6	95.6	11474	2	BD249790	Campyloba
41	837.6	95.6	11474	2	AR271699	Sequence
42	837.6	95.6	11474	2	AR481781	Sequence
43	837.6	95.6	11474	2	AR527380	Sequence
44	837.6	95.6	11474	2	AR609659	Sequence
45	837.6	95.6	11474	2	AR689937	Sequence

ALIGNMENTS

RESULT 1  
AX934431  
LOCUS AX934431 876 bp DNA linear PAT 05-JAN-2004  
DEFINITION Sequence 8 from Patent WO02074942.  
ACCESSION AX934431  
VERSION AX934431.1 GI:40641678  
KEYWORDS  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

REFERENCE 1  
AUTHORS Gilbert M. and Wakarchuk W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics  
JOURNAL Patent: WO 02074942-A 8 26-SEP-2002;  
National Research Council of Canada (CA)

FEATURES  
Location/Qualifiers  
source 1..876  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"

CDS  
1..876  
/notes="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
jejuni O:19"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04169.1"  
/db\_xref="GI:40641679"  
/translations="MKKVIAGNPSLKEIDYSLRPNDFVRCNQFEDKYVLGKK  
CKAFYTPNFFEQYTLKHLIQNETELIMCSNYNQAHLNENFKTFDYFPDA  
HGYDFPKQLKFNKFEIYFNQRTSGVNYCAVAIALGYKEIYLSGIDFYQNG  
SVAFDVKQNLKLPDKNDRSHYGHKSTDIKALEFKYKIKLYCLCPNSLLA  
NFTLAPNLNSNFIQEKNNYTKDILIPSEAYGKFSKNINFKIKIKENYVYKLIKD  
LLRLPSDIKHFKGK"

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 Db 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 QY 61 CTACCAATGATTGATGATTTAGATGTAATCAATTTTATTGAGATAAATACTAT 120  
 Db 61 CTACCAATGATTGATGATTTAGATGTAATCAATTTTATTGAGATAAATACTAT 120  
 QY 121 CTGTGTAATAAATGCAAGAGCAGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 Db 121 CTGTGTAATAAATGCAAGAGCAGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 QY 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCGAACTAATTTATGTTCTAAT 240  
 Db 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCGAACTAATTTATGTTCTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTATGATTAATTTTCT 300  
 Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTATGATTAATTTTCT 300  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 360  
 QY 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
 Db 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTC 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTC 480  
 QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTTTAAAGAAATTTAATGCTTATTAA 540  
 Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTTTAAAGAAATTTAATGCTTATTAA 540  
 QY 541 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 600  
 Db 541 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 600  
 QY 601 GAAAAAATTTAAAGAAATTTTATTTAAGAGATTTTATCAAGATTTGATTAAGATTA 660  
 Db 601 GAAAAAATTTAAAGAAATTTTATTTAAGAGATTTTATCAAGATTTGATTAAGATTA 660  
 QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATAACAGAAAAATACTACACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATAACAGAAAAATACTACACT 720  
 QY 721 AAAGATATCTATACACTTCTAGTGAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
 Db 721 AAAGATATCTATACACTTCTAGTGAGGCTTATGGAATTTTCAAAATTTTAAATTTT 780  
 QY 781 AAAAAATTTAAAGAAATTTTATTTAAGAGATTTTATCAAGATTTGATTAAGATTA 840  
 Db 781 AAAAAATTTAAAGAAATTTTATTTAAGAGATTTTATCAAGATTTGATTAAGATTA 840  
 QY 841 CCTAGTGATTAAGCAATTTTCAAGGAAATTA 876  
 Db 841 CCTAGTGATTAAGCAATTTTCAAGGAAATTA 876

RESULT 2  
 BD249794  
 LOCUS  
 DEFINITION  
 Campylobacter glycosyltransferases for biosynthesis of  
 gangliosides and ganglioside mimics.  
 ACCESSION  
 BD249794  
 VERSION  
 BD249794.1 GI:33059564  
 KEYWORDS  
 JP 2002535992-A/5.  
 SOURCE  
 Campylobacter jejuni  
 ORGANISM  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

Campylobacteraceae; Campylobacter.  
 1 (bases 1 to 876)  
 Gilbert, M. and Wakarchuk, W.W.  
 Campylobacter glycosyltransferases for biosynthesis of  
 gangliosides and ganglioside mimics  
 Patent: JP 2002535992-A 5 29-OCT-2002;  
 NATIONAL RESEARCH COUNCIL OF CANADA  
 OS Campylobacter jejuni  
 PN JP 2002535992-A/5  
 PD 29-OCT-2002  
 PF 01-FEB-2000 JP 2000597438  
 PR 01-FEB-1999 US 60/118213, 31-JAN-2000 US 06/495406 PI  
 MICHEL GILBERT, WARREN W WAKARCHUK  
 PC C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC  
 Campylobacter glycosyltransferases for biosynthesis of CC  
 gangliosides and  
 CC ganglioside mimics  
 FH Key Location/Qualifiers  
 FT source 1..876  
 FT /organism='Campylobacter jejuni'.  
 FT /mol\_type='genomic DNA'  
 FT /db\_xref='taxon:197'  
 FEATURES  
 source  
 1..876  
 /organism='Campylobacter jejuni'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:197'  
 ORIGIN  
 Query Match 100.0%; Score 876; DB 2; Length 876;  
 Best Local Similarity 100.0%; Pred. No. 9.9e-231; Indels 0; Gaps 0;  
 Matches 876; Conservative 0; Mismatches 0;  
 QY 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 Db 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 QY 61 CTACCAATGATTGATGATTTAGATGTAATCAATTTTATTGAGATAAATACTAT 120  
 Db 61 CTACCAATGATTGATGATTTAGATGTAATCAATTTTATTGAGATAAATACTAT 120  
 QY 121 CTGTGTAATAAATGCAAGAGCAGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 Db 121 CTGTGTAATAAATGCAAGAGCAGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 QY 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCGAACTAATTTATGTTCTAAT 240  
 Db 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCGAACTAATTTATGTTCTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTATGATTAATTTTCT 300  
 Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTATGATTAATTTTCT 300  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 360  
 QY 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
 Db 361 TTTTCAGGAAATTTATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTC 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAAATGGGTC 480  
 QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTTTAAAGAAATTTAATGCTTATTAA 540  
 Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTTTAAAGAAATTTAATGCTTATTAA 540  
 QY 541 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 600  
 Db 541 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTAA 600  
 QY 601 GAAAAAATTTAAAGAAATTTTATTTAAGAGATTTTATCAAGATTTGATTAAGATTA 660

Db 601 GAAAAAACTTACAAAAATAAACTATATTGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660

Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720

Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720

Qy 721 AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780

Db 721 AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780

Qy 781 AAAAAAATAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

RESULT 3

AR271703

LOCUS AR271703 876 bp DNA linear PAT 10-APR-2003

DEFINITION Sequence 8 from patent US 6503744.

ACCESSION AR271703

VERSION AR271703.1 GI:29703248

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 876)

AUTHORS Gilbert,M. and Wakarchuk,W.W.

TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics

JOURNAL Patent: US 6503744-A 8 07-JAN-2003; National Research Council of Canada; Ottawa; CAX;

FEATURES

source Location/Qualifiers

1. .876

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;

Best Local Similarity 100.0%; Pred. No. 9.9e-231;

Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTCATTATTTCAAGG 60

Db 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTCATTATTTCAAGG 60

Qy 61 CTACCAAAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

Db 61 CTACCAAAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

Qy 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

Db 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

Qy 181 ACTTTTAAACATTTTAAATCAAAATCAAGATATGACCGCAATTAATGTTGTTCTAAT 240

Db 181 ACTTTTAAACATTTTAAATCAAAATCAAGATATGACCGCAATTAATGTTGTTCTAAT 240

Qy 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATTTTACGATTATTTTCT 300

Db 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATTTTACGATTATTTTCT 300

Qy 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACCTTTAAAGAAATTTAAATCTTTATTTAA 360

Db 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACCTTTAAAGAAATTTAAATCTTTATTTAA 360

Qy 361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

Db 361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420

Qy 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480

Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 481 TCTTATGCTTTTGTATACCAACAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540

Db 481 TCTTATGCTTTTGTATACCAACAGAAATCTTTTAAACATAGCCCTGATTTTAAAAAT 540

Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAAGCTTTAGAAATTTCTA 600

Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAAAGCTTTAGAAATTTCTA 600

Qy 601 GAAAAAACTTTACAAAAATAAAACTATATTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660

Db 601 GAAAAAACTTTACAAAAATAAAACTATATTGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660

Qy 661 ATAGAACTAGCGCCAAATTTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720

Db 661 ATAGAACTAGCGCCAAATTTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720

Qy 721 AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780

Db 721 AAAGATATACTCATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780

Qy 781 AAAAAAATAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAAATAA 876

RESULT 4

AR481785

LOCUS AR481785 876 bp DNA linear PAT 14-MAY-2004

DEFINITION Sequence 8 from patent US 6699705.

ACCESSION AR481785

VERSION AR481785.1 GI:47243420

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 876)

AUTHORS Gilbert,M. and Wakarchuk,W.W.

TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics

JOURNAL Patent: US 6699705-A 8 02-MAR-2004; National Research Council of Canada; Ottawa; CAX;

FEATURES

source Location/Qualifiers

1. .876

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 876; DB 2; Length 876;

Best Local Similarity 100.0%; Pred. No. 9.9e-231;

Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60

Db 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60

Qy 61 CTACCAAAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

Db 61 CTACCAAAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120

Qy 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

Db 121 CTGTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180



REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert,M. and Wakarchuk,W.W.  
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity  
JOURNAL Patent: US 682019-A 8 30-NOV-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source Location/Qualifiers  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
DB 1 ATCAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGCACTAATATGCTTCTTAAT 240  
DB 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGCACTAATATGCTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTAATGCTTATTTAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTAATGCTTATTTAA 360  
QY 361 TTTTCAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
DB 361 TTTTCAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAA 540  
DB 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAA 540  
QY 541 GATCGCTCGCATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAATATAATTCCTATGTCCTTAATAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTTACAAATATAATTCCTATGTCCTTAATAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAAATACTACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAAATACTACACT 720  
QY 721 AAAGATATACATACCTTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTT 780  
DB 721 AAAGATATACATACCTTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTT 780  
QY 781 AAAAAATATAAATAATTAAGAAATGTTTATTAAGAGTTGATAAAGCTTTAGAAATTTA 840  
DB 781 AAAAAATATAAATAATTAAGAAATGTTTATTAAGAGTTGATAAAGCTTTAGAAATTTA 840  
QY 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAAAATAA 876

Db 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAAAATAA 876  
RESULT 7  
LOCUS AR689941 876 bp DNA linear PAT 12-SEP-2005  
DEFINITION Sequence 8 from patent US 6905867.  
ACCESSION AR689941  
VERSION AR689941.1 GI:74471949  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert,M. and Wakarchuk,W.W.  
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity  
JOURNAL Patent: US 6905867-A 8 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source Location/Qualifiers  
1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 876; DB 2; Length 876;  
Best Local Similarity 100.0%; Pred. No. 9.9e-231;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
DB 1 ATCAAAAAGTTATTATTCCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGCACTAATATGCTTCTTAAT 240  
DB 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGCACTAATATGCTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAA 360  
QY 361 TTTTCAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
DB 361 TTTTCAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAA 540  
DB 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAA 540  
QY 541 GATCGCTCGCATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTACAAATATAATTCCTATGTCCTTAATAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTTACAAATATAATTCCTATGTCCTTAATAGTCTTTTACCAATTTT 660

```
Db 601 GAAAAAAGCTTACAAATAAAGCTATATTGCTTATGCTCTAATAGTCTTTTACCAAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAAATACTACACT 720
Qy 721 AAAGATATACATACCTTTAGTGGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATACATACCTTTAGTGGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Qy 781 AAAAAATATAAATTAAGAAATCTTTATACAAAGTTGATTAAGAGCTATTAGATTA 840
Db 781 AAAAAATATAAATTAAGAAATCTTTATACAAAGTTGATTAAGAGCTATTAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAATAA 876

RESULT 8
LOCUS AR691837 876 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 8 from patent US 6911337.
ACCESSION AR691837
VERSION AR691837.1 GI:74479548
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAc transferase
JOURNAL Patent: US 6911337-A 8 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 876; DB 2; Length 876;
Best Local Similarity 100.0%; Pred. No. 9.9e-231;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAGAAATTCATTTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGCTTTAAAGAAATTCATTTCAAGG 60
Qy 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
Qy 121 CTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCGGAATTAATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCGGAATTAATGTTCTTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATCAAAATTTTGAAGAACTTTTACGATTTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAATCAAAATTTTGAAGAACTTTTACGATTTTCTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTA 360
Qy 361 TTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420
Db 361 TTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420
```

```
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTCATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAAATACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAAAAATACTACACT 720
Qy 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTT 780
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTT 780
Qy 781 AAAAAAATAAATAATTAAGAAATGTTTATTACAAAGTTGATTAAGAGCTATTAGATTA 840
Db 781 AAAAAAATAAATAATTAAGAAATGTTTATTACAAAGTTGATTAAGAGCTATTAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAATAA 876
Db 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAATAA 876

RESULT 9
LOCUS AR661458 6047 bp DNA linear BCT 20-JAN-2006
DEFINITION Campylobacter jejuni strain CF90-26 truncated putative glucosyltransferase, putative glycosyltransferase, beta-1,4-N-acetylglucosaminyltransferase (cgta), beta-1,3-galactosyltransferase (cgtB), and alpha-2,3-sialyltransferase (cstII) genes, complete cds; and sialic acid synthase (neuB) gene, partial cds.
ACCESSION AR661458
VERSION AR661458.1 GI:56159787
KEYWORDS
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
REFERENCE 1 (bases 1 to 6047)
AUTHORS Koga, M., Gilbert, M., Takahashi, M., Li, J., Koike, S., Hirata, K. and Yuki, N.
TITLE Comprehensive Analysis of Bacterial Risk Factors for the Development of Guillain-Barre Syndrome after Campylobacter jejuni Enteritis
JOURNAL J. Infect. Dis. 193 (4), 547-555 (2006).
PUBMED 16425134
REFERENCE 2 (bases 1 to 6047)
AUTHORS Koga, M., Gilbert, M., Karwaski, M. F., Leclerc, S. and Yuki, N.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2004) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada
FEATURES
source Location/Qualifiers
1..6047
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="CF90-26"
/serotype="HS:19"
/db_xref="taxon:197"
57..1313
/note="Orf1"
/codon_start=1
CDS
```





```

DB      5711 CCTAGTGATATAAGCATTATTTCAGAGGAAATAA 5746
|||||
RESULT 10
LOCUS   AY422197
DEFINITION
Campylobacter jejuni putative ABC-type transport protein, putative
UDP-glucose-4-epimerase, hypothetical protein, putative
heptosyltransferase I, putative lipid A biosynthesis
acyltransferase, putative two-domain glycosyltransferase, putative
glycosyltransferase, putative
beta-1,4-N-acetylgalactosaminyltransferase, putative
beta-1,3-galactosyltransferase, putative
alpha-2,3-sialyltransferase, putative
putative UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine
synthase, putative CMP-NeuAc synthetase, putative
acetyltransferase, putative glycosyltransferase, putative
heptosyltransferase II, putative phosphoheptose isomerase 1,
putative ADP-heptose synthase, putative
ADP-glyceromannoheptose-6-epimerase, putative phosphatase, putative
periplasmic cytochrome C, small hypothetical hydrophobic protein,
and putative cation-transporting ATPase genes, complete cds; and
putative transcription termination factor gene, partial cds.
AY422197
AY422197.1 GI:40217896
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Campylobacteraceae; Campylobacter.
1 (bases 1 to 24425)
Li, J., Wakarchuk, W.W. and Endtz, H.P.
Evidence for Acquisition of the Lipopoligosaccharide Biosynthesis
Locus in Campylobacter jejuni GB11, a Strain Isolated from a
Patient with Guillain-Barre Syndrome, by Horizontal Exchange
Infect. Immun. 72 (2), 1162-1165 (2004)
14742567
2 (bases 1 to 24425)
Gilbert, M., Karwaski, M.-F., Cunningham, A. and Brochu, D.
Direct Submission
Submitted (24-SEP-2003) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
source
Location/Qualifiers
1..24425
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="GB11"
/serotype="HS:2"
/db_xref="taxon:197"
complement (21..1715)
/codon_start=1
/transl_table=11
/product="putative ABC-type transport protein"
/protein_id="AAR82866.1"
/db_xref="GI:40217897"
/translation="MLKKLFFILSKEDKNLFFLLVFSVFISFIETFAISLVMPITLL
ASDFSYDRNKLISLEYDNIIPVEIIVFVGLVIFVFRALNAYPHLLARFSK
GRYHAIAYKFSKFINVEKFTQKNQEILKSTIGEYVNLSTMISSFLLMSEIFV
LLVYALMLINKYITFLSIFWLNAPILVKLSPIKKAGVRREAMKNFELIINTV
LNNEFKLTKEDGVLSEFKQAQSEAFSKANITNESVAAPRIVLEGIGFCVLFIV
FLVKNESDISGILSTISFVLYALRYLMPSANRITNESYHDLLYHSSLDIYQNRQE
BENIGEEKLSQBELKICNLSFGVEKKYLPKNLNLKKEKIAFISGCGKSTLV
DLI1GLKPKGQQLIDQELNANNNTKRYKIGYIPQNIYLFNDSIARNTFFGDAV
BEKNRLYKQANLHFIFKLPQGVOTKVGDSNLSGGOKRIARALYLEPEMLVL
DEATNSALDQSEAKIMDIYKISKDKXTMIIIAHRLSTITQCDKVYRLEHGLKEEK"
complement (1709..2695)
/codon_start=1
/transl_table=11
/product="putative UDP-glucose-4-epimerase"
/protein_id="AAR82867.1"
/db_xref="GI:40217898"
/translation="MKILISGGAGYIGSHTLROFLKTDHIEICVLDNLKSGSKIAIEDL
OKIRAFKFEODLSDFQGVKALPEREKFDALVHFAASIEVFESMONPLKYVMNNTVNT
TNLIETCLOTGVNKFIFSSAATYGPQTPVSETSLAPINPYGRSKLMSSEVLRLDA
SWANPEFKCIILRYFNVAAGACMDYTLGQRYPKATLLIKVAAEACAGRKDKLFIKDDY
DKDGTICIDFIHVDDIISAHALDYLKENESNVFNVGHGFSVKKEVIEAMKKVSG
VDFKVELAPRRAGDPSVLISDASKIRNLTSWQPKYDDELICKSAFDEWKQC"
complement (2753..3583)
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAR82868.1"
/db_xref="GI:40217899"
/translation="MILAKFLKKLWKAKNEGYICVDFCESVPDVELIRKTLGPEGSD
LEVSLKALQWKEQSGSEFLPLPVHKIISTCAVSDNFGKFIVKNIDGONEKEMIEIN
FNFNIENYKPLVSENGKMFDMPLVLRALKYNLKAATYLDLTQSDKWNKYKTRSELK
HDLLESLSNGRGIKDLTCLSCNVGLPGKYDVHGDEVKMLFYENELKIKIYBCSDVL
NTMULFKYELIKANVSEBDYVDFLSMRDFLRKKSDRSYTEVFAKACESEISKVS"
3613..4641
/codon_start=1
/transl_table=11
/product="putative heptosyltransferase I"
/protein_id="AAR82869.1"
/db_xref="GI:40217900"
/translation="MKIAIVRLKALGDIIOQSAVVLOFIFKNFKKDIIEHWFVDEKFEGI
LKNHPLIDKLYALPLKDKKILKSLKILLKARKNNYNAVILQGLIKSAIVSRILSRNN
FGFDKNSLSEFAHNFYNQKLELDYNNVFRVLSLTSFMLNTDFNAKNLAFKQDVS
VBNLQKLNKLLDKNEKNILIHVSSVENKIYPATKLAILLCLLIDNFGPQTKIWL
ANGVNEYFAFELVNLGIDETHEIELAPKFNLELMFAKTMMDLLIIGNDSGGTHLAF
ALNKASITITFGATPSYRNAFQTHINKIIDTGKKIQNAKHIDKSDFCITRIEEDIFKL
AKGLNEK"
4631..5518
/codon_start=1
/transl_table=11
/product="putative lipid A biosynthesis acyltransferase"
/protein_id="AAR82870.1"
/db_xref="GI:40217901"
/translation="MKNSDRIVLSYLVILKFFVTMPDCTILHFLALIVARTAPHLNKK
HRKILNTNIOICFPQYQTKERDKLSKIYENFAQGDICLQNTQTEKILNKYNFIN
ENFLDALAKRPIIPTAHYGNWEILSLAYAAKYGVISVGRKLKSEVMEIILSQSR
TQFDIELDKKGIROMLSALKKRALGILTDQDCVENESVRLKFNKFNKYNQOASL
IAQRNNAIIPVYAYKEGKFCIEFFKAKDSQNASLEELTYLQAQSCCEMITKKRPWEY
FFPHRRFASYNEEIYVGAK"
5515..6768
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAR82871.1"
/db_xref="GI:40217902"
/translation="MNLKQISVIIIVKNAEQTLLECLNSLKDFDEIILLNNESSDNTLL
KIANEFKDQFANLYIYHNAFIFGKALKNLASYAKNDWILSDADEVLENECIELKN
LKQEDNIILSRKNLYKGEWIKACGWPDYVLRIFKNKTFRENDNLVHESLVLPSNA
KKIYDNLGLKHSYKDIISHLIDMQYYSLLMAKQNTHKKSVLKNLARAFWTFPNTF
KNGFLYGVKGFIIISVCSALGTFFYKMLYLOKQPKTCALLITYNQERKLVLVDS
VKNLAFLPNEVLIADDGSKEDTARLIEEYKQFPCLKHIWODEGFGKLSKRNTIK
NADSEYIIVIDGMILEKDFIKHLEFAQRKLLQGSRVILNKKSEIILNKDORYII
FNKKDFKSSNFKLAKIFYLSKKR"
7064..8233
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAR82872.1"
/db_xref="GI:40217903"
/translation="MKKIGWIPYINVEKYLRECLDSVINQTYTNLEIILVNDGSTDE
HSLNTAKETLADKRITLFDKNGKGIASSARNIGIEYFSGYKLNKTQHKENSLEIF
QLDGNPNYIKAYKSSQAPNNEKDLNTFTYPSIDIIFLSDNYWKLNCIEECVIRM
KNVDVLDHDCCTYEDNIKNKHKTEMEIFDFKECIIITPEYANRLKSVGSRDISPC
WNGMDIENFLKQIKLKFIPINEDIHFGIILPASANKIYVLSQLYLCILRANSIN
HDKITKANVSEYFKDIYETFGENAKEKNYLKAASRVITALLKLIIBFFKQOKNNAIA
IKETFLPCYAKKAMIKKFKOPNLNLEQLVLIKPFIQTQTKLPYDIWKFWQIKNI"
8303..9346
/codon_start=1

```



PUBMED 10660542  
REFERENCE 2 (bases 1 to 24437)  
AUTHORS Gilbert, M., Cunningham, A.-M., Karwaski, M.-F., Michniewicz, J.,  
Wu, Y., Young, N.M. and Wakarchuk, W.W.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUL-1999) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada  
REFERENCE 3 (bases 1 to 24437)  
AUTHORS Gilbert, M.  
TITLE Direct Submission  
JOURNAL Submitted (22-JAN-2004) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada  
REMARK Nucleotide and amino acid sequences updated by submitter  
COMMENT On Jan 22, 2004 this sequence version replaced gi:6969981.  
FEATURES  
source  
1. .24437  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="ATCC 43446"  
/serotype="HS.19"  
/db\_xref="ATCC:43446"  
/db\_xref="taxon:197"  
complement (21. .1715)  
/codon\_start=1  
/transl\_table=11  
/product="putative ABC-type transport protein"  
/protein\_id="AAR99161.1"  
/db\_xref="GI:41058384"  
/translation="MLKLLFLLSKEDKNFLFFLLVFSVFSFIETFAISLVMPFITL  
ASDFYDRNKLISLEYLNPVFEIIVFVGVLIVFVFRALINAYFFLLARFSK  
RHHAIAYKVSFKNINVEKFTQKNOSEILKTSVYNLSTWISSLLMSEIFTV  
LLVALLMLNKKITLFLSIFMVLNATLILKIPKILAGLREEMAKNEFIILTN  
LNPFKTLKTKEGVSLFKAQSEAFSKANITNESVAAVRILEGFGCVLFIIV  
FLVNKESDLSIGLUSTISFVIALYRLMPSANRITSHVDLLLYHSSLDIYIQNLRQ  
EENIGELKSNQELKCNLSFGYEGKHLFNLNINLKKEKIAFTEGSCGKSTLV  
DLIIGLKEGGQILIDKQELNASNARKYKIGYIPONIYLFNDSIAKNTFFGDAV  
BEKLSKVIKOANLHFTKNLPGQVTKVGGGSLNGQKQKORIAARALYLEPFTLV  
DEATSAIDTQSEAKIMDEIYKISKDKTMIIAHRLSITITQCDKYVRLHGLKEEK"  
complement (1709. .2695)  
/codon\_start=1  
/transl\_table=11  
/product="putative UDP-glucose-4-epimerase"  
/protein\_id="AAR99162.1"  
/db\_xref="GI:41058385"  
/translation="MKLISGAGYIGSHTRLQFLKTDHETICVLDNLSKGSKIALEDL  
QKRAFPEODLSDFQGVKALFEREKDAIVHAASIEVPESQNPPLKYINNVTNT  
TNLIETCMQTVNRFIFSSATYGEPTVPVSETSPAPINPYGRSKLMEELVLRDA  
SMANPEFKHILRYFNAGACMDYTLGQRYPKATLLIKVAAECAGKRDKLFIQDDY  
DTKDGTCTIRDFIHVDDISSAHALADYLOENESVFNNGVGHGFSVKEVIAEMKVS  
VDFKVELAPRAGDPSVLSADSKIRNLTSQPKYDDLLELICKSAFDEKQC"  
complement (2753. .3583)  
/inference="non-experimental evidence, no additional  
details recorded"  
/codon\_start=1  
/transl\_table=11  
/product="hypothetical protein"  
/protein\_id="AAR99163.1"  
/db\_xref="GI:41058386"  
/translation="MILAKLKKRLKMAKNEGIVCFDCEVPDPVELIRKTLIGPEGSD  
LEVSLKALQKQESGEFLPLPVHKIISICAVLSDNFGFKIKYKNDIGQNEKEMIN  
FFNFIENVEPKLVGKNGKNDMPVLVLPALAKYKATLYLDTQSKNNYKTRFSELK  
HCDLLELSGNGRIKLDITLCSMGLPGKYDVHGDVWMLFYENELEKIYIYCSDDL  
NTNWLFLYELIKANVEEDYVDFLSYMRDFLRKKSDRSYTEVFAKACESEISKVR  
S"  
3613. .4641  
/note="formerly ORF-01; similar to Helicobacter pylori  
RfaC"  
/codon\_start=1  
/transl\_table=11  
/product="heptosyltransferase I"  
/protein\_id="AAF34138.2"  
/db\_xref="GI:41058382"  
/translation="MKIAIVRLSALGDIISQAVVLOFIKFNPKDIEIHFVDFEKEFEG  
LKNHPLDKLVALPLKDKLSKILLKARKNNYNAVLDQGLIKSAIVSRILSRNN  
FDGKNSLKESPAHNFYNOKLELDYNNENFVRYLSLTSMNTDFNAKNALFKQDVFS  
VDENLKOLLNNKLDKNEKNLIIHVGSSEVENKIYPTKLAAILCKLILNFSQOQKIWL  
AWGNVKEVFAKEVNLSDIGETIELAPFNDEELWAFATKMDMLLIIGNSGPHTLAF  
AKLKASITIFGATPSYRNAFQTHINKIIDTGKKIQNAKHIDKSDFCITRIEEDIFKL  
AKGLLNEK"  
4631. .5518  
/note="formerly ORF-02; similar to Helicobacter pylori  
WaaM"  
/codon\_start=1  
/transl\_table=11  
/product="acylttransferase"  
/protein\_id="AAF34139.1"  
/db\_xref="GI:6969984"  
/translation="MKNSDRIYLSLYILKFFVTFMPDCLLHFLALIVARIAFLNKK  
HRKIINTNLQICPQYTKERDKLSKIYENFAQFGIDCLQNTTKERILKNVNFIN  
ENFLDALALKEPIIFTTAHYGNWEILSLAYAAKYGAISTVKGKLSVMEYILSQSR  
TQFIDELDKGGIRQMSALKKERALGILTDQDCVENESVRLKFFNKVYQMGASL  
IAORSNALIIPVAVYKEGGKCTIEFFKAKDSQNASLEELTYQAQSCSEMIKKRPWEY  
PEFRRFASYNEEIYKGAK"  
5515. .6771  
/note="formerly ORF-03; similar to Neisseria meningitidis  
LgtF"  
/codon\_start=1  
/transl\_table=11  
/product="glycosyltransferase"  
/protein\_id="AAF34140.1"  
/db\_xref="GI:6969985"  
/translation="MNLKQISVIIIVKNAEOTLLECLNSLKDFEILLNNESDNTL  
KIANEFKDFANLIYHNAGIFGALKNLALSAYAKNDWILSIDADEVLENECKELKN  
KLQSDNITALLSKNLYKGEWIKACQWPPDYLRIFNKNFTRENDNLVHESLVLPNSA  
KKIILKNGLKHSYKDIKSLIDKMOXYSSIMAKNTHKKSGLKANLARAFWTFPRNYE  
LKNGLYKVGKFIISVCNAGLTFKMKLYELQRPKTCALIIITYNQKERLKLVD  
SVKNLAFIPNEVLIIADDSKEDTARLIEEYQKDFPCPLKHIMWOEDGFKLSKGRKNTI  
KNADSEYIIVIDGMILEKDFIKHELFQAQRLFQGSRVILNKKESEETLNKDDYRI  
IFNKKDFKSSFLAKIFYSLSKKR"  
7067. .8236  
/note="formerly ORF-04; similar to Streptococcus  
pneumoniae Cps14J"  
/codon\_start=1  
/transl\_table=11  
/product="glycosyltransferase"  
/protein\_id="AAF34141.1"  
/db\_xref="GI:6969986"  
/translation="MKKIGVPIPIVNEKVLRECLDSVINTQYTNLEIILVNDGSTDE  
HSLNIAKSYTLKDKRITLFDKNGGLSSARNIGIYFSGEYKLKNNKTOHKENSLEIF  
QLDGNPNYNIYKAYSQAFNNEKDLTNFTYPSIDYIIIFLSDNYWKLNGICECVMIR  
KNVDVLPFDHDCITYEDNINKHKTRMEIFDFKKECIITPKYANRALSVGRDISFG  
WNGMIDENFLKQIKLKFINTFINEDIHFGIILFASANKIYVLKQYLCLRLANSIN  
HDKKTITKANVSEYFKDIYETFGENAKEAKNVLKASRVITALKLIEFFKDKQENALA  
IKETFLPCYAKKALMIKPKFDPLNLEQVLVLPKIPQITKLPDYIMKFWOKIKNI"  
8306. .9349  
/gene="cgtA"  
8306. .9349  
/gene="cgtA"  
/note="similar to Campylobacter jejuni OH4384 CgtA encoded  
by GenBank Accession Number AF130984"  
/codon\_start=1  
/transl\_table=11  
/product="glycosyltransferase"  
/protein\_id="AAF34142.1"  
/db\_xref="GI:6969987"  
/translation="MLFOSYFVKIICLFIPIFRKIRHKIKKTFLLKNIQRDKIDSYLPK  
KTLQINKYNNEDLILKNAKIRREGHKGFYNDKSKDPKSPINWAFIRVNEAITL  
KASLSYLPALQORGIVGDGDEEILEFCQKQYSPFIPKYPVEIQIQNPKSEEN  
KLSYSYNVVAFIPKDEWLIKIDVDHIVDAKLYKSFYPKNYDVVVSVDIHYFN  
DNFLCKDNNDNLKEPGDCLLNNYLNKWEVLIDRINNKKATQKQSFSSHIHSL  
QLKYKHLRIFHTLNNYHFFFLKGRQAQDIYKYNWISIEEFKFKYIQNINHKIPESMI  
SKETLKKIFLTLF"  
complement (9335. .10240)  
/note="cgtB; similar to Campylobacter jejuni OH4384 CgtB

```
encoded by GenBank Accession Number AF130984"
/codon_start=1
/transl_table=11
/product="glycosyltransferase"
/protein_id="AAF34143.1"
/db_xref="GI:6969988"
/translation="MFKISILILPTYNVEOYIARIESCINOTFKDIEIIVDPCGNDN
SINIAKEYSKDKRIKIIHNEKLGILLRAREGVKVANSPIINFLPDDDDVLELNACEE
CKIKDEQDELIVFNAIVESNVISYKPFDFNSGYFSKKEFVKKIIAKGLWYTWG
KUIRKLXLEAFLEKDEKINMAEDVLLYPMLSQAKIAYMNCNLHYHVPNNNS
ICNTKNEVLKNNIOBLQVLNRYLNRYILNKYCSVLYLVLYLQIYKIKRTKLM
VTLLAKINILTLKILPKYKFLKQC"
10329..11204
/gene="cst-II"
10329..11204
/gene="cst-II"
/experiment="experimental evidence, no additional details
recorded"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAF34143.1"
/db_xref="GI:6969988"
/translation="MKKVIILAGNPSLKEIDYSRLEPDEFVFCNQFYPEDKYLGKK
CKAVYTPNPFPEQYVTLKHLIQNEYETELIMCSNNOAHLENFVKTFYDFDA
HUGYDFKOLKEFNAYFKPEHIFNORITSGVYMCVAIALGYKEIKLYLGSIDFYONGS
SYAFPTQENLLKLAIDPFKDRSHVIGHSKNTDIIKALEFLEKTYKIKLYGLCPNSLLA
NFIELAPNLNSNFIIOEKNNYTKDILIPSEAYGKPSKNINPKKIKIKENVYKLIKID
LRLPLSDIKHYFKGK"
11205..12245
/note="formerly ORF-08; similar to Neisseria meningitidis
SiaC"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAF34144.1"

Query Match
Best Local Similarity 100.0%; Score 876; DB 15; Length 24437;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAGATTATTATCTCGGAATGACCAAGTTTAAAGAAATGATATTCAAGG 60
Db 10329 ATGAAAAGATTATTATCTCGGAATGACCAAGTTTAAAGAAATGATATTCAAGG 10388
QY 61 CTACCAATGATTTGATGATTAATGATGATTAATCAATTTTATTTCAAGATAAATACTAT 120
Db 10389 CTACCAATGATTTGATGATTAATGATGATTAATCAATTTTATTTGAAGATAAATACTAT 10448
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTCTTGAGCAATACTAC 180
Db 10449 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTTCTTGAGCAATACTAC 10508
QY 181 ACTTTAAACATTTAATCAAAATCAAGAAATGACCGCACTAAATATGTTCTTAAT 240
Db 10509 ACTTTAAACATTTAATCAAAATCAAGAAATGACCGCACTAAATATGTTCTTAAT 10568
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATATTTCCT 300
Db 10569 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATATTTCCT 10628
QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360
Db 10629 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 10688
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTCAGTAGCC 420
Db 10689 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTCAGTAGCC 10748
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAATGGGTCA 480
Db 10749 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATGATTTTATCAAAATGGGTCA 10808
```

```
QY 481 TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAAGACTAGCCCTGATTTTAAAAAT 540
Db 10809 TCATTATGCTTTTGATACCAACCAAGAAATCTTTTAAAGACTAGCCCTGATTTTAAAAAT 10868
QY 541 GATCGCTCGCATATATCGGCACATAGTAAATAATACAGATATATAAAGCTTTAGAAATTTCTA 600
Db 10869 GATCGCTCGCATATATCGGCACATAGTAAATAATACAGATATATAAAGCTTTAGAAATTTCTA 10928
QY 601 GAAAAAACTTACAAAAATAAACTATATTTGCTTTATGCTCTCTTAATAGTCTTTTAGCAAAATTTT 660
Db 10929 GAAAAAACTTACAAAAATAAACTATATTTGCTTTATGCTCTCTTAATAGTCTTTTAGCAAAATTTT 10988
QY 661 ATGAACTAGCGCCCAAAATTTAAATTTCAATTTTATCATAACAAGAAAAATAAATACATACACT 720
Db 10989 ATGAACTAGCGCCCAAAATTTAAATTTCAATTTTATCATAACAAGAAAAATAAATACATACACT 11048
QY 721 AAGATATATCTCATACCTCTCTAGTGAGCTTATGAGAAATTTTCAAAAAATAATTAATTTT 780
Db 11049 AAGATATATCTCATACCTCTCTAGTGAGCTTATGAGAAATTTTCAAAAAATAATTAATTTT 11108
QY 781 AAAAAATAAAAAATAAAGAAATTTTATTTACAAAGTTGATAAAGAGATCTTATTAAGATTA 840
Db 11109 AAAAAATAAAAAATAAAGAAATTTTATTTACAAAGTTGATAAAGAGATCTTATTAAGATTA 11168
QY 841 CCTAGTATATAAAGCAATTTATTTCAAGGAAAAATAA 876
Db 11169 CCTAGTATATAAAGCAATTTATTTCAAGGAAAAATAA 11204

RESULT 12
AF305571 4749 bp DNA linear BCT 29-JAN-2002
LOCUS Campylobacter jejuni strain 81-176 sialyl transferase (cst), sialic
DEFINITION acid synthase (neuB), NeuC1, N-acetylgalactosyltransferase, and
CMF-NeuNAc synthetase (neuA1) genes, complete cds.
ACCESSION AF305571
VERSION AF305571.1 GI:15824490
KEYWORDS Campylobacter jejuni subsp. jejuni 81-176
SOURCE Campylobacter jejuni subsp. jejuni 81-176
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 4749)
AUTHORS Guerry, P., Szymanski, C.M., Prendergast, M.M., Hickey, T.E.,
Ewing, C.P., Pattarini, D.L. and Moran, A.P.
TITLE Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide
affects ganglioside mimicry and invasiveness in vitro
JOURNAL Infect. Immun. 70 (2), 787-793 (2002)
PUBMED 11796612
REFERENCE 2 (bases 1 to 4749)
AUTHORS Guerry, P., Szymanski, C.M., Prendergast, M.M., Hickey, T.E.,
Ewing, C.P., Pattarini, D.L. and Moran, A.P.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2000) Enteric Diseases, Naval Medical Research
Center, 503 Robert Grant Ave, Silver Spring, MD 20910, USA
FEATURES
source
1..4749
/organism="Campylobacter jejuni subsp. jejuni 81-176"
/mol_type="genomic DNA"
/strain="81-176"
/db_xref="taxon:354242"
89..964
/gene="cst"
89..964
/gene="cst"
/notes="cst"
/codon_start=1
/transl_table=11
/product="sialyl transferase"
/protein_id="AA109368.1"
/db_xref="GI:15824491"
/translation="MKKVIILAGNPSLKEIDYSRLEPDEFVFCNQFYPEDKYLGKK
CKAVYTPNPFPEQYVTLKHLIQNEYETELIMCSNNOAHLENFVKTFYDFPDA
```

HLGVDFPKQLKEFNAYKFKHEIYFNQRIYTSYVYMCVAVIAIGYKEIYLSGIDFYQNGS  
SYAFDTQENLLKAPDFKNDRSHYIGHSXNTDIKALFLEKTYKIKYCLCPNSLLA  
NFIELAPLNSNFIIQKNNYTKDILIPSSAYGFKSKNINFKKIKIKENYVYKLIKID  
LLRLPSDIKHYFKGK"

gene

/gene="neuB1"

CDS

/gene="neuB1"

/note="neuB1"

/codon\_start=1

/transl\_table=11

/product="sialic acid synthase"

/protein\_id="AAL09369.1"

/db\_xref="GI:15824492"

GAKIHKTHIVEDMSKAAKVIPGNAKISYEIMOKALDYKDELAKYKTEKLG  
VLIPTSPRAGANLEDMGVSAFKIGSGCENYPLIKHIAAFKPMIYSTGMSIESI  
KPTVKILLDNEIPVLMHSTNLYPTPHNLVRLNMLELKKFSCMVGUSDHTDNLAC  
LGAVALGACVLERHFTDSMRSGPDIVCSMDTKALKELIIQSEQMAIMKGNESKAA  
KQEQVTIDFAPASVVSINDIKKGEVLSMDNIWVRPGLGGISAAEFENILGKKALRDI  
ENDTQLSYEDFA"

CDS

/note="N-acetylmannosamine synthesis"

/codon\_start=1

/transl\_table=11

/product="NeuC1"

/protein\_id="AAL09370.1"

/db\_xref="GI:15824493"

/translation="MYRVONSSEPELYIFATGMHLSKNFQVTVKELYKNGKPNYIEFI  
NDKYQDKALATIDGFSRYNELKEDLIVVHGDRLEPLAAIIVGALNINLVAHIE  
GGESITDLSRLAIKSLAHILVNDDEFKRLMOLGEDEKSPFIQSPDELLNDN  
KISLEAKYDINVENYALIMHPVTEITSINKQADNLVKALIQSNKNTIVIYPPN  
DILGPELILQSVBELKNNPRFKLPSPLRPEYITLLKNADFIIGNSSCILKEALYKTA  
GILVGRONGRLNGENTLKVANSDEILKAINTHKKQDLFSTKLEILDSSKLFPEYL  
LSGFFPKLSTQKVKDIKHA"

CDS

/note="CgtA"

/codon\_start=1

/transl\_table=11

/product="N-acetylglactosyltransferase"

/protein\_id="AAL09371.1"

/db\_xref="GI:15824494"

/translation="MLKKIISLYKRYISKSLVDNEHFIFIKENKNIYKKGKHPDFD  
EKADKVPKPMWGFIRVKNELTLRVLSLESLPALORGIAYNDCDGSBELILEFC  
KOPNFYAKPYVDLENKPNENKLYSYNWAASPIPLDEWFIKIDVDHYDAKKL  
YKSVRIDQENKALCYPRINFIILNGNIYVQNSGYPIGGDQLLIKERNSEFLER  
VSKSQMIDPKGLIEEYSEQQVLSQGVKILQAPLLQWHFFALKTRNDYQQYLDILS  
LEEFQAFHRKSKEAKKIDFAMLKRPIVEQILKKFQGEIK"

gene

/gene="neuA1"

CDS

/gene="neuA1"

/note="neuA1"

/codon\_start=1

/transl\_table=11

/product="CMP-NeuNAC synthetase"

/protein\_id="AAL09372.1"

/db\_xref="GI:15824495"

/translation="MSLAILPARGSGKIKKNNVLNNKPLIYYTIKAALNAKISK  
VVSDDSEILNAYKSNQVNDILKRPISLAQDDTSDKVLHLALFKYDYDVFVLOPT  
SPLRNTHIDKAFNLKYNNSANALISVSECKNLKAFVCDNDGLAGICNDYFPMP  
ROKLPKTYMNSGAIYILKIKFELNNPSPLQNKTHFLWDESSSLDIDCLDLKKVEQI  
WKK"

ORIGIN

Query Match 99.8%; Score 874.4; DB 15; Length 4749;

Best Local Similarity 99.9%; Pred. No. 1.3e-230;

Matches 875; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGCTTTTAAAGAATAATGATTATTCAGG 60

|||||

89 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGCTTTTAAAGAATAATGATTATTCAGG 148

|||||

Qy 61 CTACCAAAATGATTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAAATACAT 120  
Db 149 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAAATACAT 208  
Qy 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATAC 180  
Db 209 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATAC 268  
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAAATATATGTGTCTTAAT 240  
Db 269 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAAATATATGTGTCTTAAT 328  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCCCT 300  
Db 329 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCCCT 388  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 389 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 448  
Qy 361 TTTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAGTAGCC 420  
Db 449 TTTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCAGTAGCC 508  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATAGGGTCA 480  
Db 509 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTATCAAAATAGGGTCA 568  
Qy 481 TCTTATGCTTTTATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 569 TCTTATGCTTTTATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 628  
Qy 541 GATCGCTCGCATCTATCGGACATAGCTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600  
Db 629 GATCGCTCACATCTATCGGACATAGCTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 688  
Qy 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Db 689 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 748  
Qy 661 ATAGAACTAGCCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAATAATACACT 720  
Db 749 ATAGAACTAGCCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAATAATACACT 808  
Qy 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 809 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 868  
Qy 781 AAAAAAATAAAAAATTAAGAAAATGTTTATTACAACTTGATAAAAAGATCTTATTAGATTA 840  
Db 869 AAAAAAATAAAAAATTAAGAAAATGTTTATTACAACTTGATAAAAAGATCTTATTAGATTA 928  
Qy 841 CCTAGTGATATTAAGCAATTTTTCAGGAAATAA 876  
Db 929 CCTAGTGATATTAAGCAATTTTTCAGGAAATAA 964

RESULT 13

AF401528

LOCUS

DEFINITION

AF401528 12390 bp DNA linear BCT 03-JAN-2002  
Campylobacter jejuni O:36 heptosyltransferase I (waaC) gene,  
partial cds; lipid A biosynthesis acyltransferase (htrB), putative  
two-domain glycosyltransferase, putative glycosyltransferase,  
truncated beta-1,4-N-acetylglactosaminyltransferase (cgtA-I),  
beta-1,3-galactosyltransferase (cgtB), alpha-2,3-sialyltransferase  
(cst-II), sialic acid synthase (neuB1), putative  
N-acetylglucosamine-6-phosphate 2-epimerase (neuC1),  
beta-1,4-N-acetylglactosaminyltransferase (cgtA-II), CMP-Neu5Ac  
synthetase (neuA1), putative acetyltransferase, and putative  
heptosyltransferase (waaV) genes, complete cds; and  
heptosyltransferase II (waaF) gene, partial cds.

ACCESSION

AF401528

VERSION

AF401528.1 GI:15718482

KEYWORDS  
SOURCE  
ORGANISM

Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

REFERENCE

1 (bases 1 to 12390)  
Gilbert, M., Karwaski, M.F., Bernatchez, S., Young, N.M., Taboada, E.,  
Michniewicz, J., Cunningham, A.M., and Wakarchuk, W.W.  
The Genetic Bases for the Variation in the Lipid-oligosaccharide of  
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF  
SIALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE  
J. Biol. Chem. 277 (1), 327-337 (2002)

JOURNAL  
PUBMED

REFERENCE

2 (bases 1 to 12390)  
Gilbert, M., Michniewicz, J., Karwaski, M.-F., Cunningham, A. and  
Wakarchuk, W.W.  
Direct Submission

TITLE  
JOURNAL

Submitted (22-JUL-2001) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive., Ottawa, Ontario K1A  
0R6, Canada

FEATURES  
source

1. .12390  
Location/Qualifiers  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="ATCC 43456"  
/serotype="O:36"  
/db\_xref="ATCC:43456"  
/db\_xref="taxon:197"  
<1..348  
/gene="waaC"  
/gene="waaC"  
<1..348  
/gene="waaC"  
/codon\_start=1  
/transl\_table=11  
/product="heptosyltransferase I"  
/protein\_id="AAL05984.1"  
/db\_xref="GI:15718483"

gene

CDS

translation="FAKEVLNLSGIDETHIELAPKFNLEELMAFTKMDLIIGNDGGP  
THLAPALNKAISITIFGATPSYRNAQTINKIIDTGKKIQNAKIDKSDFCITRIEE  
DIFLAKGLNEK"  
338..1225  
/gene="htrB"  
338..1225  
/gene="htrB"  
/codon\_start=1  
/transl\_table=11  
/product="lipid A biosynthesis acyltransferase"  
/protein\_id="AAL05985.1"  
/db\_xref="GI:15718484"

gene

CDS

translation="MKNSDRIYLSLYILKFFVTWMPDCILHFLALIVARIAFLNKK  
HRKIINTLQICFPQYTKQKDKLSKIVENFAQFIDCLQONQTTKEKILNKNVFIN  
ENFLDALAKRPILFTTAHYNWELISLAVAAKGAISIVCKLSEVWVILQSQR  
TQFDLELDKGGIROMLSALKERALGILTDQCVNESVRKFTFKREVNTQMGASL  
IAQRNALIIVYAKKEGKFCIEFFKARDSONASLEBLTYOQSCCEMIKGRPMWEY  
FFFHRRFASNEYIKGAK"  
1222..2769  
/note="orf3"  
/codon\_start=1  
/transl\_table=11

gene

CDS

product="putative two-domain glycosyltransferase"  
/protein\_id="AAL05986.1"  
/db\_xref="GI:15718485"  
translation="MNLKQISVIIIVKNAEOTLLECLNSLKDFDEIILLNPSIDNTL  
NIANFEKDFANLYIYHNAFIGFGALKNLALSYAKNDWILSIDADEVLENCIKELKN  
LELEDQNIILSRNLYKGEWIKACGWSPDYVIRFNKFNTRFNLDNLVHESGIVLPSNV  
KKIHLKGLHYAFRDISHLINKQWYSILWAKQNHKKSGVILNKLRAFWTFPFRFY  
LVNGLYGYQGFITSVNCALGTFYKMLYELQKQKPTCALIITTYNOKERLKLVLVD  
SVKNLTFLPNEVLADGSKEDTARLIEYOKDPCSLKHWOEDEGFAAKNRNNAI  
RVSKSEYIILLDGMILEKDFIKNHFEKAKRVLILQGSVILNKNSEILKNDQDYSL  
AFNKQDFKSYNKLAKIYIKISKIDQKFPNTKEILQSKTCNMSSFRAKDFDELQGFN  
ENFTGWDHEDSEFVARFLFNKGIIFRLRKFAIAHYHYEENSKMLESNHNQIYLDTIK  
NKKISWR"

gene

CDS

translation="MKKVIAGNPSLKEIDYRLPNDPFRNQCFYFEDKYYIGKK  
CKTFTYTPNPFQYTYTHLQIQEYETELIMCSNYNQAHLENENFVKTDPYDPDA  
HLYGDFPKQLKEFNAPFKFHEIYFNQRTSGVMCAVAIALGYKEIYLSGIDFYQNGS  
SVAFDTQENLLKLAPDFKNDRSHYIGHSKNTDIKALEFLEKTKYKLYCLCPNSLLA  
NFIELAPNLNSFILOEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKD  
LURLPSDIKHFFKQK"  
6886..7926  
/gene="neuB1"  
6886..7926  
/gene="neuB1"  
/codon\_start=1  
/transl\_table=11

gene

CDS

product="sialic acid synthase"  
/protein\_id="AAL05991.1"  
/db\_xref="GI:15718490"  
translation="MKEIKQNTIISSEKAPLVVPEIGINHNGSLAKIMVDAFSA  
GAKIKHQTHIVEDENSKAKKVIIPGNAKISIEYIMQKCALDYDELAKKEYTEKLG  
VYLTSTPSRAGANRFLSDMGVSAPFKIGSGECNNYPLIKHIAAFKPKMIVSTGMNSES  
KPTKTLIDLENEIPFLMTNTNPTPHNLVRLNMLLEKFEFSQMGVLSLSDHNSIESI  
LGAVALGACVLERHFTDSMRSGPDIVCSMDTKALKELIIQSEQAMTKMGNESKAA  
KQEQVTIDFAFASVWSIKDIKGEVLSMDNIWVKRPGLGISAAEFENILGKALRDI

gene

CDS

translation="MKEIKQNTIISSEKAPLVVPEIGINHNGSLAKIMVDAFSA  
GAKIKHQTHIVEDENSKAKKVIIPGNAKISIEYIMQKCALDYDELAKKEYTEKLG  
VYLTSTPSRAGANRFLSDMGVSAPFKIGSGECNNYPLIKHIAAFKPKMIVSTGMNSES  
KPTKTLIDLENEIPFLMTNTNPTPHNLVRLNMLLEKFEFSQMGVLSLSDHNSIESI  
LGAVALGACVLERHFTDSMRSGPDIVCSMDTKALKELIIQSEQAMTKMGNESKAA  
KQEQVTIDFAFASVWSIKDIKGEVLSMDNIWVKRPGLGISAAEFENILGKALRDI

gene

CDS

translation="MKEIKQNTIISSEKAPLVVPEIGINHNGSLAKIMVDAFSA  
GAKIKHQTHIVEDENSKAKKVIIPGNAKISIEYIMQKCALDYDELAKKEYTEKLG  
VYLTSTPSRAGANRFLSDMGVSAPFKIGSGECNNYPLIKHIAAFKPKMIVSTGMNSES  
KPTKTLIDLENEIPFLMTNTNPTPHNLVRLNMLLEKFEFSQMGVLSLSDHNSIESI  
LGAVALGACVLERHFTDSMRSGPDIVCSMDTKALKELIIQSEQAMTKMGNESKAA  
KQEQVTIDFAFASVWSIKDIKGEVLSMDNIWVKRPGLGISAAEFENILGKALRDI

gene

CDS

translation="MKEIKQNTIISSEKAPLVVPEIGINHNGSLAKIMVDAFSA  
GAKIKHQTHIVEDENSKAKKVIIPGNAKISIEYIMQKCALDYDELAKKEYTEKLG  
VYLTSTPSRAGANRFLSDMGVSAPFKIGSGECNNYPLIKHIAAFKPKMIVSTGMNSES  
KPTKTLIDLENEIPFLMTNTNPTPHNLVRLNMLLEKFEFSQMGVLSLSDHNSIESI  
LGAVALGACVLERHFTDSMRSGPDIVCSMDTKALKELIIQSEQAMTKMGNESKAA  
KQEQVTIDFAFASVWSIKDIKGEVLSMDNIWVKRPGLGISAAEFENILGKALRDI

/note="orf4"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAL05987.1"  
/db\_xref="GI:15718486"  
translation="WEIKHEENRCSYTNIQRRRIERFMFNSVINQTYTNLEILLVNDG  
STDENLNIAREYTLRDFILDKNGGLSSARNVGLIEYFSGEYKJKNKTHIKENS  
LIEFOLDGNPNYIKAYKQSAFNENKEDLTNTFYNDIYIFLSDDDYWEJLNCIEEC  
VIRMKNDVLFWDHDCYEDNKNKHKKTRMEIFDFKKECIITPEYANRALSVSGRD  
ISFGWGMIDFNLKIKLAFINFINEDIHFGIILPASANKIYILSKLYLCRLRAN  
SISNHDKAITKANVSEYFKDIYETFFGENAKENYLAASRTVTKALKLIEFFPKQKSE  
NALAIKETFLPCAKGALMIKKFKKQFLNLKQLVLKPFITQKLPIYDIWKFWQKIKK  
IN"  
3988..4248  
/gene="cgtA-I"  
3988..4248  
/gene="cgtA-I"  
/note="orf5-I; premature translation stop due to a one bp  
deletion"  
/codon\_start=1  
/transl\_table=11  
/product="truncated  
beta-1,4-N-acetylgalactosaminyltransferase"  
/protein\_id="AAL05988.1"  
/db\_xref="GI:15718487"  
translation="MLFQSYFVKIICLFIPFRKIRHKIKTKTFLKKNIQDKIDSYLK  
KILIQINKYNNEDLIKLNKAITGGGIKIDILMMKNLKIQNLL"  
complement (5016..5921)  
/gene="cgtB"  
complement (5016..5921)  
/gene="cgtB"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,3-galactosyltransferase"  
/protein\_id="AAL05989.1"  
/db\_xref="GI:15718488"  
translation="MFKISILPTVYVQYIARIESCINOTFKDIEIIVVDDCGNDN  
SINAKESKKDKRIKIHNKGLLRLARVEGVKANSPIVIMFLDPDDYLELNACEE  
KILDEODEVLAVFNAIVEVSIVSKYKDFDNGSGFYSKKEPVKKIATKNLYWTWNG  
KLIRKKLYLEAFASRLRKDVKNMAEDVLIIYPMLSQAQKIAMNCLNLYHYVPNNNS  
ICNTKNEVLVNNIIOELVLNLRQYILNLYKYSVLVLIYLLYIYQIYKIKRTKLM  
VTLLAKINITLTKILFKYKFLKQC"  
6010..6885  
/gene="cst-II"  
6010..6885  
/gene="cst-II"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-2,3-sialyltransferase"  
/protein\_id="AAL05990.1"  
/db\_xref="GI:15718489"  
translation="MKKVIAGNPSLKEIDYRLPNDPFRNQCFYFEDKYYIGKK  
CKTFTYTPNPFQYTYTHLQIQEYETELIMCSNYNQAHLENENFVKTDPYDPDA  
HLYGDFPKQLKEFNAPFKFHEIYFNQRTSGVMCAVAIALGYKEIYLSGIDFYQNGS  
SVAFDTQENLLKLAPDFKNDRSHYIGHSKNTDIKALEFLEKTKYKLYCLCPNSLLA  
NFIELAPNLNSFILOEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKD  
LURLPSDIKHFFKQK"  
6886..7926  
/gene="neuB1"  
6886..7926  
/gene="neuB1"  
/codon\_start=1  
/transl\_table=11  
/product="sialic acid synthase"  
/protein\_id="AAL05991.1"  
/db\_xref="GI:15718490"  
translation="MKEIKQNTIISSEKAPLVVPEIGINHNGSLAKIMVDAFSA  
GAKIKHQTHIVEDENSKAKKVIIPGNAKISIEYIMQKCALDYDELAKKEYTEKLG  
VYLTSTPSRAGANRFLSDMGVSAPFKIGSGECNNYPLIKHIAAFKPKMIVSTGMNSES  
KPTKTLIDLENEIPFLMTNTNPTPHNLVRLNMLLEKFEFSQMGVLSLSDHNSIESI  
LGAVALGACVLERHFTDSMRSGPDIVCSMDTKALKELIIQSEQAMTKMGNESKAA  
KQEQVTIDFAFASVWSIKDIKGEVLSMDNIWVKRPGLGISAAEFENILGKALRDI

ENDTQLSVEDFA"  
7923. .9047  
/gene="neuc1"  
7923. .9047  
/gene="neuc1"  
/codon\_start=1  
/transl\_table=11  
/product="putative N-acetylglucosamine-6-phosphate  
2-epimerase"  
/protein\_id="AAL05992.1"  
/db\_xref="GI:15718491"  
/translation="MKKILFITGRADYSKISLMYRVQNSSEFELYIPATGMHLKSN  
FGYTVKLYKNGFNKNIYEFINVDKYQTDKALATTIDGFSRYVNEKLPDLVWHDRI  
EPLAAIVGALNNILVAHIEGGEISGTTDDSRHAISKLAHILVNDVFAKRRMLQD  
EDEKSIPIIGSPDLLELNDKISLSEAKKYVDINVENYVALLMFHPTVTEITSIKNOAD  
NLKVALIQSNKNIVYVNNDLGPDELILQSVLEELKNRPFKLPSLRFEYFETILKNA  
DFIIGNSCILKEALYILKTAGILVGSRRNGRLGNENTILKVNANSDEILKAINTHKKQ  
DLFSTKLELDSSKLPFFELYSGEFFVLSLSTQKVFKDIIKHA"  
9040. .9987  
/gene="cgtA-II"  
9040. .9987  
/gene="cgtA-II"  
/note="orf5-II"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,4-N-acetylglucosaminyltransferase"  
/protein\_id="AAL05993.1"  
/db\_xref="GI:15718492"  
/translation="MLKKILSYKRSISKKLVLDNEHFIEKNKIYKSKHGPFDF  
EKADKSPNPMWGFIRVKNALTLRVSLISLPALQRIIAYNDCDDGSBELILEFC  
KQYFNFAIDKPYVDLENPKNEKNLYSYVNWASFLPDWFKIDVDHYDAKLL  
YKSFYRIDQENKALCYPRINFILNGNIYVQNSNGYFIGGDDOLLIKRRNSSPTERR  
VSKSQWIDPKGLIEELYSEQOVLQOAPLLOWHPALVKRRNDYQOYLDILS  
LEEFQAFHRKSKEAKKIDFAMLRPVIVQILKKPQGEIK"  
9984. .10649  
/gene="neuA1"  
9984. .10649  
/gene="neuA1"  
/codon\_start=1  
/transl\_table=11

Query Match 99.6%; Score 872.8; DB 15; Length 12390;  
Best Local Similarity 99.8%; Pred. No. 2.4e-230;  
Matches 874; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 ATGAAAAGTATTATCTGGAATCGACCAAGTTTAAAGAAATGATATTCAAG 60  
DB 6010 ATGAAAAGTATTATCTGGAATCGACCAAGTTTAAAGAAATGATATTCAAG 6069  
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120  
DB 6070 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTGAAGATAAATACTAT 6129  
QY 121 CTGCTAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 180  
DB 6130 CTGCTAAAAATGCAAGCAGTGTTTACACCCCTAATTTCTCTTTGAGCAATACTAC 6189  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGACCGCAACTAATTTATGTTCTAAT 240  
DB 6190 ACTTTAAACATTTAATCCAAATCAAGAAATATGACCGCAACTAATTTATGTTCTAAT 6249  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 300  
DB 6250 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 6309  
QY 301 GATGCTCATTTGGGATGATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 360  
DB 6310 GATGCTCATTTGGGATGATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAA 6369  
QY 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 420  
DB 6370 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGTCAGTAGCC 6429

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTTTATCAAAATGGGTCA 480  
DB 6430 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTTTATCAAAATGGGTCA 6489  
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
DB 6490 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6549  
QY 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600  
DB 6550 GATCGCTCACATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 6609  
QY 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 660  
DB 6610 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 6669  
QY 661 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACACT 720  
DB 6670 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACACT 6729  
QY 721 AAAGATATATCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTTAAATTT 780  
DB 6730 AAAGATATATCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTTAAATTT 6789  
QY 781 AAAAAAATAAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
DB 6790 AAAAAAATAAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 6849  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 876  
DB 6850 CCTAGTGATATAAGCAATTTTCAAGGAAATAA 6885  
RESULT 14  
AX934436  
LOCUS  
DEFINITION  
AX934436  
ACCESSION  
AX934436.1 GI:40641682  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE  
1 Gilbert, M. and Wakarchuk, W.W.  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
Patent: WO 02074942-A 13 26-SEP-2002;  
National Research Council of Canada (CA)  
Location/Qualifiers  
1. .873  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1. .>873  
/note="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
jejuni O:36"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04171.1"  
/db\_xref="GI:40641683"  
/translation="MKKVIAGNPSLKEIDYRLPNDPFDVFCNQFYFEDKYYLGGK  
KLGTYFTNPFEEQYITLKLHINOYEETELINCYNQAHENENFVKTFYDFDA  
HGYDFQKLFENAYFKFHEIYFNQITSGVYMCVAVALGYEIKYLSGIDFYQNGS  
SVAPTKOENLKLAPDFKNDKSHVIGHKNTDIIKALEFKYKIKLYCLCPNLSLA  
NFIELAPNLNSFIQEKNNYTKDILIPSEAYGKFSKINFKIKIKENVYKLIKID  
LLRLPSDIKHFKGK"  
ORIGIN  
Query Match 99.3%; Score 869.8; DB 2; Length 873;  
Best Local Similarity 99.8%; Pred. No. 5.1e-229;



Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	ATGAAAAAGTTATATTCTGGAATGACCAAGTTTAAAGAAATGATGATTTCAAGG	60
DB	1	ATGAAAAAGTTATATTCTGGAATGACCAAGTTTAAAGAAATGATGATTTCAAGG	60
QY	61	CTACCAATGATTTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT	120
DB	61	CTACCAATGATTTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT	120
QY	121	CTTGTTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC	180
DB	121	CTTGTTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC	180
QY	181	ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAATAATTAATGTTCTTAAT	240
DB	181	ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAATAATTAATGTTCTTAAT	240
QY	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTAATTTTCT	300
DB	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTAATTTTCT	300
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA	360
DB	301	GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA	360
QY	361	TTTCACGAAATTTAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC	420
DB	361	TTTCACGAAATTTAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC	420
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
QY	481	TCCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT	540
DB	481	TCCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT	540
QY	541	GATCGCTCAGCTATATCGGATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA	600
DB	541	GATCGCTCAGCTATATCGGATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA	600
QY	601	GAAAAAATCAAAATTAATGCTTATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAAT	660
DB	601	GAAAAAATCAAAATTAATGCTTATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAAT	660
QY	661	ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATACTACACT	720
DB	661	ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATACTACACT	720
QY	721	AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT	780
DB	721	AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT	780
QY	781	AAAAAATTAATAATTAAGAAATGTTTATTAAGAGTCTAATAAGATTA	840
DB	781	AAAAAATTAATAATTAAGAAATGTTTATTAAGAGTCTAATAAGATTA	840
QY	841	CCTAGTGATATAAGCATTTTCAAGGAAA	873
DB	841	CCTAGTGATATAAGCATTTTCAAGGAAA	873

RESULT 15  
AR481787  
LOCUS AR481787  
DEFINITION Sequence 13 from patent US 6699705.  
ACCESSION AR481787  
VERSION AR481787.1 GI:47243422  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 873) AUTHORS Gilbert, W. and Wakarchuk, W. W. TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics JOURNAL Patent: US 6699705-A 13 02-MAR-2004; National Research Council of Canada; Ottawa; CAX;			
FEATURES source 1..873 /organism="unknown" /mol_type="genomic DNA"			
ORIGIN			
Query Match 99.3%; Score 869.8; DB 2; Length 873; Best Local Similarity 99.8%; Pred. No. 5.1e-229; Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	ATGAAAAAGTTATATTCTGGAATGACCAAGTTTAAAGAAATGATGATTTCAAGG	60
DB	1	ATGAAAAAGTTATATTCTGGAATGACCAAGTTTAAAGAAATGATGATTTCAAGG	60
QY	61	CTACCAATGATTTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT	120
DB	61	CTACCAATGATTTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT	120
QY	121	CTTGTTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC	180
DB	121	CTTGTTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC	180
QY	181	ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAATAATTAATGTTCTTAAT	240
DB	181	ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAATAATTAATGTTCTTAAT	240
QY	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTAATTTTCT	300
DB	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTAATTTTCT	300
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA	360
DB	301	GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAA	360
QY	361	TTTCACGAAATTTAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC	420
DB	361	TTTCACGAAATTTAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGCC	420
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	480
QY	481	TCCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT	540
DB	481	TCCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT	540
QY	541	GATCGCTCAGCTATATCGGATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA	600
DB	541	GATCGCTCAGCTATATCGGATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA	600
QY	601	GAAAAAATCAAAATTAATGCTTATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAAT	660
DB	601	GAAAAAATCAAAATTAATGCTTATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAAT	660
QY	661	ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATACTACACT	720
DB	661	ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATACTACACT	720
QY	721	AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT	780
DB	721	AAAGATATACCTATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT	780
QY	781	AAAAAATTAATAATTAAGAAATGTTTATTAAGAGTCTAATAAGATTA	840
DB	781	AAAAAATTAATAATTAAGAAATGTTTATTAAGAGTCTAATAAGATTA	840



Qy 841 CCTAGTGATATATAAGCATTATTTCAAAGGAAAA 873  
| | | | | | | | | | | | | | | | | | | | | |  
Db 841 CCTAGTGATATATAAGCATTATTTCAAAGGAAAA 873

Search completed: July 19, 2006, 19:12:30  
Job time : 5069.9 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:02:20 ; Search time 136.645 Seconds  
(without alignments)  
11995.226 Million cell updates/sec

Title: US-10-734-719-8  
Perfect score: 876  
Sequence: 1 atgaaaaaagtattattgc.....attatttcaaggaataaa 876

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 93554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1 COMB.seq:\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5 COMB.seq:\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A COMB.seq:\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B COMB.seq:\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7 COMB.seq:\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H COMB.seq:\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PCTUS COMB.seq:\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP COMB.seq:\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE COMB.seq:\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles!.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	876	100.0	876	3	US-09-495-406-8
2	876	100.0	876	3	US-09-816-028A-8
3	876	100.0	876	3	US-10-303-162-8
4	876	100.0	876	3	US-10-303-134-8
5	876	100.0	876	3	US-10-303-118-8
6	876	100.0	876	3	US-10-303-128-8
7	876	100.0	876	5	US-10-735-419-8
8	869.8	99.3	873	3	US-09-816-028A-13
9	869.8	99.3	873	3	US-10-303-162-13
10	869.8	99.3	873	3	US-10-303-134-13
11	869.8	99.3	873	3	US-10-303-118-13
12	869.8	99.3	873	3	US-10-303-128-13
13	869.8	99.3	873	5	US-10-735-419-13
14	866.6	98.9	873	3	US-09-816-028A-11
15	866.6	98.9	873	3	US-10-303-162-11
16	866.6	98.9	873	3	US-10-303-134-11
17	866.6	98.9	873	3	US-10-303-118-11
18	866.6	98.9	873	3	US-10-303-128-11
19	866.6	98.9	873	5	US-09-495-406-4
20	844	96.3	876	3	US-09-816-028A-4
21	844	96.3	876	3	US-10-303-162-4
22	844	96.3	876	3	US-10-303-134-4
23	844	96.3	876	3	US-10-303-118-4

24	844	96.3	876	3	US-10-303-118-4	Sequence 4, Appli
25	844	96.3	876	3	US-10-303-128-4	Sequence 4, Appli
26	844	96.3	876	5	US-10-735-419-4	Sequence 4, Appli
27	837.6	95.6	11474	3	US-09-495-406-1	Sequence 1, Appli
28	837.6	95.6	11474	3	US-09-816-028A-1	Sequence 1, Appli
29	837.6	95.6	11474	3	US-10-303-162-1	Sequence 1, Appli
30	837.6	95.6	11474	3	US-10-303-134-1	Sequence 1, Appli
31	837.6	95.6	11474	3	US-10-303-118-1	Sequence 1, Appli
32	837.6	95.6	11474	3	US-10-303-128-1	Sequence 1, Appli
33	837.6	95.6	11474	5	US-10-735-419-1	Sequence 1, Appli
34	836	95.4	876	3	US-09-495-406-2	Sequence 2, Appli
35	836	95.4	876	3	US-09-816-028A-2	Sequence 2, Appli
36	836	95.4	876	3	US-10-303-162-2	Sequence 2, Appli
37	836	95.4	876	3	US-10-303-134-2	Sequence 2, Appli
38	836	95.4	876	3	US-10-303-118-2	Sequence 2, Appli
39	836	95.4	876	3	US-10-303-128-2	Sequence 2, Appli
40	836	95.4	876	5	US-10-735-419-2	Sequence 2, Appli
41	820	93.6	876	3	US-09-495-406-6	Sequence 6, Appli
42	820	93.6	876	3	US-09-816-028A-6	Sequence 6, Appli
43	820	93.6	876	3	US-10-303-162-6	Sequence 6, Appli
44	820	93.6	876	3	US-10-303-134-6	Sequence 6, Appli
45	820	93.6	876	3	US-10-303-118-6	Sequence 6, Appli

## ALIGNMENTS

### RESULT 1

US-09-495-406-8  
; Sequence 8, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495\_406  
; CURRENT FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni O:19

Query Match 100.0%; Score 876; DB 3; Length 876;  
Best Local Similarity 100.0%; Pred. No. 5.7e-160; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAAAAAAGTTATTATTCGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60
Db	1	ATGAAAAAAGTTATTATTCGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60
Qy	61	CTACCAAAATGATTTTGATGTATTAGATGAATCAATTTTATTTGAAGATAAATCTAT	120
Db	61	CTACCAAAATGATTTTGATGTATTAGATGAATCAATTTTATTTGAAGATAAATCTAT	120
Qy	121	CTTGGTAAAAAATGCAAGCAGTCTTTTACACCCCTAATTTCTTTGAGCAATCTAC	180
Db	121	CTTGGTAAAAAATGCAAGCAGTCTTTTACACCCCTAATTTCTTTGAGCAATCTAC	180
Qy	181	ACTTTAAAAACATTAATCCAAATATGAGACCGCAACTAATTTGTGTTCTAAT	240

181	Db	ACITTTAAACATTTTAATCCAAAATCAGAAATATGAGACCGAACTAATTAATGTTGTTCTAAAT	240
241	Qy	TACAACCAAGCTCATCTAGAAAAATGAAAAAATTTGTAAAAAATCTTTTACGAAATATTTTCCT	300
241	Db	TACAACCAAGCTCATCTAGAAAAATGAAAAAATTTGTAAAAAATCTTTTACGAAATATTTTCCT	300
301	Qy	GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAAAAATTTTAAAGAAATTTTAAAGTCTTATTTTAA	360
301	Db	GATGCTCATTTGGGATATGATTTTAAAAAATTTTAAAAAATTTTAAAGAAATTTTAAAGTCTTATTTTAA	360
361	Qy	TTTTCACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGGCC	420
361	Db	TTTTCACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGGCC	420
421	Qy	ATAGCCCTAGGATACAAAGAAATTTATCTCTTCGCGGAATTGATTTTTATCAAAATCGGGTCA	480
421	Db	ATAGCCCTAGGATACAAAGAAATTTATCTCTTCGCGGAATTGATTTTTATCAAAATCGGGTCA	480
481	Qy	TCATTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAAAATAGCCCTTGATTTTAAAAAAT	540
481	Db	TCATTATGCTTTTGATACCAAAACAGAAAAATCTTTTAAAAATAGCCCTTGATTTTAAAAAAT	540
541	Qy	GATCGCTCGCATATATATCGGCATAGTAAAAATACAGATATAAAAAAGCTTTAGAAATTTCTA	600
541	Db	GATCGCTCGCATATATATCGGCATAGTAAAAATACAGATATAAAAAAGCTTTAGAAATTTCTA	600
601	Qy	GAAAAAATTTACAAAAATAAAACTATATGTCTTATGTCTCTTAATAGTCTTTTACGAAATTTT	660
601	Db	GAAAAAATTTACAAAAATAAAACTATATGTCTTATGTCTCTTAATAGTCTTTTACGAAATTTT	660
661	Qy	ATAGAACTAGCGCCAAAATTTTAAAATTTCAAATTTTATCATACAAGAAAAAAAATAACTACACT	720
661	Db	ATAGAACTAGCGCCAAAATTTTAAAATTTCAAATTTTATCATACAAGAAAAAAAATAACTACACT	720
721	Qy	AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT	780
721	Db	AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT	780
781	Qy	AAAAAAATAAAAATTTAAAGAAAAATGTTTTATTACAAGTTGATAAAAAAGATCTATTAAGATT	840
781	Db	AAAAAAATAAAAATTTAAAGAAAAATGTTTTATTACAAGTTGATAAAAAAGATCTATTAAGATT	840
841	Qy	CCTAGTGATATAAAGCATATTTTCAAAGGAAAAATAA	876
841	Db	CCTAGTGATATAAAGCATATTTTCAAAGGAAAAATAA	876

## RESULT 2

```

RESULI' 2
US-09-816-028A-8
; Sequence 8, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; . APPLICANT: Gilbert, Michel
; . APPLICANT: Wakarchuk, Warren W.
; . APPLICANT: National Research Council of Canada
; . TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; . TITLE OF INVENTION: Campylobacter Gangliosides and Ganglioside Mimics
; . FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; . FILE REFERENCE: 019633-00011UUS
; . CURRENT APPLICATION NUMBER: US/09/816,028A
; . CURRENT FILING DATE: 2001-03-21
; . PRIOR APPLICATION NUMBER: US 60/118,213
; . PRIOR FILING DATE: 1999-02-01
; . PRIOR APPLICATION NUMBER: US 09/495,406
; . PRIOR FILING DATE: 2000-01-31
; . NUMBER OF SEQ ID NOS: 49
; . SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)

```

GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-162-8

Query Match 100.0%; Score 876; DB 3; Length 876;  
Best Local Similarity 100.0%; Pred. No. 5,7e-160; Indels 0; Gaps 0;  
Matches 876; Conservative 0; Mismatches 0;

QY 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAGG 60

QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACAT 120  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACAT 120

QY 121 CTTGTGTAATAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTGAGCAATAC 180  
DB 121 CTTGTGTAATAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTGAGCAATAC 180

QY 181 ACTTTAAACATTTTAAATCAAGATATGAGACCGAACTAAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTTAAATCAAGATATGAGACCGAACTAAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACCGATTATTTGCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACCGATTATTTGCT 300

QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAA 360  
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAA 360

QY 361 TTTTCAGGAATTTATTTCATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTACC 420  
DB 361 TTTTCAGGAATTTATTTCATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTACC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCCTCGCCTATATCGACATAGTAAATATACAGATATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCCTCGCCTATATCGACATAGTAAATATACAGATATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660

DB 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACACT 720  
QY 721 AAAGATATACCTACCTCTAGTGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATTAATTT 780  
DB 721 AAAGATATACCTACCTCTAGTGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATTAATTT 780  
QY 781 AAAGATATACCTACCTCTAGTGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATTAATTT 840  
DB 781 AAAGATATACCTACCTCTAGTGGCTTATGAGGCTTATGAGGCTTATGAGGCTTATTAATTT 840  
QY 841 CCTAGTATATATAAGCATTATTTCAAAGGAAATAA 876  
DB 841 CCTAGTATATATAAGCATTATTTCAAAGGAAATAA 876

RESULT 4  
US-10-303-134-8  
; Sequence 8, Application US/10303134  
; Patent No. 6825019  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-134-8

Query Match 100.0%; Score 876; DB 3; Length 876;  
Best Local Similarity 100.0%; Pred. No. 5,7e-160; Indels 0; Gaps 0;  
Matches 876; Conservative 0; Mismatches 0;

QY 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTAATGCTTATTTCAAG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATTAATGCTTATTTCAAG 60

QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACAT 120  
DB 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACAT 120

QY 121 CTTGTGTAATAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTGAGCAATAC 180  
DB 121 CTTGTGTAATAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTCTTGAGCAATAC 180

QY 181 ACTTTAAACATTTTAAATCAAGAAATTAATGCTTATGAGCAGCACTAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTTAAATCAAGAAATTAATGCTTATGAGCAGCACTAATATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACCGATTATTTGCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACCGATTATTTGCT 300

QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAA 360  
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAA 360

QY 361 TTTTCAGGAATTTATTTCATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTACC 420  
DB 361 TTTTCAGGAATTTATTTCATCAAGAAATTTACCTCAGGGTCTATATGTTGCGAGTACC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCCTCGCCTATATCGACATAGTAAATATACAGATATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCCTCGCCTATATCGACATAGTAAATATACAGATATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAACTTACAAAATAAACTATATGCTTATGCTTAAATAGTCTTTTAGCAATTTT 660

```

Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACATTTTACGATTATTTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAAATGCTTTATTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAAATGCTTTATTTAA 360
Qy 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGC 420
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGGC 420
Qy 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 480
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 480
Qy 481 TCTTATGCTTTTATACCAACCAAGAAATTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTATACCAACCAAGAAATTTTAAACTAGCCCTGATTTTAAAT 540
Qy 541 GATGCTCGCATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCT 600
Db 541 GATGCTCGCATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCT 600
Qy 601 GAAACCTTACAAATAAACTATATTTGCTTATGCTTCTTAACTAGCTTTTACCAATTTT 660
Db 601 GAAACCTTACAAATAAACTATATTTGCTTATGCTTCTTAACTAGCTTTTACCAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATAAGAAATTTTCAAAATTTTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATAAGAAATTTTCAAAATTTTACACT 720
Qy 721 AAAGATATATCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
Db 721 AAAGATATATCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
Qy 781 AAAAAATAAATTTAAAGAAATTTTATTTACAGTTGATAAAAGATCTATTAGATTA 840
Db 781 AAAAAATAAATTTAAAGAAATTTTATTTACAGTTGATAAAAGATCTATTAGATTA 840
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876

```

```

RESULT 5
US-10-303-118-8
; Sequence 8, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2.3/alpha 2.8-sialyltransferase II

```

```

; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-118-8
Query Match 100.0%; Score 876; DB 3; Length 876;
Best Local Similarity 100.0%; Pred. No. 5.7e-160;
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTATTATTTGCTGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Db 1 ATGAAAAAGTATTATTTGCTGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60
Qy 61 CTACCAAAATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAAAATGATTTTATGATGTTATAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Qy 121 CTTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATCTAC 180
Db 121 CTTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATCTAC 180
Qy 181 ACTTTAAAAATTAATCCAAATCAAGAAATATGAGACCGAACTTAATTTATGTGTCTAAT 240
Db 181 ACTTTAAAAATTAATCCAAATCAAGAAATATGAGACCGAACTTAATTTATGTGTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTTAAA 360
Qy 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGTACC 420
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGTACC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTTCATACCAACCAAGAAATTTTAAACTAGCCCTGATTTTAAAT 540
Db 481 TCTTATGCTTTTTCATACCAACCAAGAAATTTTAAACTAGCCCTGATTTTAAAT 540
Qy 541 GATGCTCGCATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCT 600
Db 541 GATGCTCGCATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCT 600
Qy 601 GAAAAAATTTACAAATAAATCTATATTTGCTTATGCTTATGCTTATGCTTATGCTTATG 660
Db 601 GAAAAAATTTACAAATAAATCTATATTTGCTTATGCTTATGCTTATGCTTATGCTTATG 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATAAGAAATTTTCAAAATTTTAAATTTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATAAGAAATTTTCAAAATTTTAAATTTT 720
Qy 721 AAAGATATATCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
Db 721 AAAGATATATCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
Qy 781 AAAAAATAAATTTAAAGAAATTTGTTTATTTACAGTTGATAAAAGATCTATTAGATTA 840
Db 781 AAAAAATAAATTTAAAGAAATTTGTTTATTTACAGTTGATAAAAGATCTATTAGATTA 840
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAATAA 876

```

```

RESULT 6
US-10-303-128-8
; Sequence 8, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:

```

APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/303,128  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US/09/816,028  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-128-8

Query Match 100.0%; Score 876; DB 3; Length 876;  
Best Local Similarity 100.0%; Pred. No. 5.7e-160;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAAAATGTTTATGATGCTATGATGCTAAATCAATTTTATTTGAGATAATATCTAT 120  
DB 61 CTACCAAAATGTTTATGATGCTATGATGCTAAATCAATTTTATTTGAGATAATATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGAGAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATATCTAC 180  
DB 121 CTGTGTAATAAATGCAAGAGAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATATCTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTATTTTAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTAATGCTTATTTTAA 360  
QY 361 TTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420  
DB 361 TTTCAGGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
QY 541 GATCGCTCGCATATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCGCATATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATCTTACAAATAAATACTATATGCTTATGCTTAAATGCTTTTACGAAATTTT 660

DB 601 GAAAAAATCTTACAAATAAATACTATATGCTTATGCTTAAATGCTTTTACGAAATTTT 660  
QY 661 ATAGAACTAGCGCCCAAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720  
DB 661 ATAGAACTAGCGCCCAAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720  
QY 721 AAAGATATACATCATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTT 780  
DB 721 AAAGATATACATCATACCTTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTT 780  
QY 781 AAAAAAATAAAAATTTAAAGAAAAATCTTTTATTAACAAGTTTGTATAAAGATCTATTAAAGATTA 840  
DB 781 AAAAAAATAAAAATTTAAAGAAAAATCTTTTATTAACAAGTTTGTATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 876  
DB 841 CCTAGTGATATAAAGCAATTTTCAAAGGAAAAATAA 876

RESULT 7  
US-10-735-419-8  
Sequence 8, Application US/10735419  
Patent No. 7026147  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-000111US  
CURRENT APPLICATION NUMBER: US/10/735,419  
CURRENT FILING DATE: 2003-12-11  
PRIOR APPLICATION NUMBER: US/09/816,028A  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-735-419-8

Query Match 100.0%; Score 876; DB 5; Length 876;  
Best Local Similarity 100.0%; Pred. No. 5.7e-160;  
Matches 876; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTTGCTGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATTTGCTGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60  
QY 61 CTACCAAAATGATTTTATGATGATTTAGATGATTAATCAATTTTATTTGAGAGATAAATCTAT 120  
DB 61 CTACCAAAATGATTTTATGATGATTTAGATGATTAATCAATTTTATTTGAGAGATAAATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGAGAGTGTGTTTACACCCCTAAATTTCTTTTGGAGCAATATCTAC 180  
DB 121 CTGTGTAATAAATGCAAGAGAGTGTGTTTACACCCCTAAATTTCTTTTGGAGCAATATCTAC 180  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTTAAT 240  
DB 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAATTAATATGTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCT 300

Db 241 TACACCAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACAGATTATTTCTCT 300  
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATCTTATTTTAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATCTTATTTTAA 360  
Qy 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCCTCGCATATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTTAAAGTTCTA 600  
Db 541 GATCCTCGCATATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTTAAAGTTCTA 600  
Qy 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660  
Db 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720  
Db 661 ATAGAACTAGCCCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720  
Qy 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAA 780  
Db 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAA 780  
Qy 781 AAAAAATTTAAATTTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
Db 781 AAAAAATTTAAATTTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 876

## RESULT 8

US-09-816-028A-13  
; Sequence 13, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: Gangliosides and Ganglioside Mimics  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1998-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-09-816-028A-13

Query Match 99.3%; Score 869.8; DB 3; Length 873;  
Best Local Similarity 99.8%; Pred. No. 8.9e-159;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTATTTGCTGGAATGGAACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTTGCTGGAATGGAACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Qy 61 CTACCAAAATGATTTTGTATGATGATTAATCAATTTTATTTTGAAGATATAATCTAT 120  
Db 61 CTACCAAAATGATTTTGTATGATGATTAATCAATTTTATTTTGAAGATATAATCTAT 120  
Qy 121 CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATCTAC 180  
Db 121 CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTAGCAATCTAC 180  
Qy 181 ACTTTAAAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 240  
Db 181 ACTTTAAAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAATTTTACGATATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAATTTTACGATATTTTCT 300  
Qy 301 GATGCTCAATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAAATTTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAAATTTTAAA 360  
Qy 361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAAGTTCTA 600  
Db 541 GATCGCTCGCATATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAAGTTCTA 600  
Qy 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660  
Db 601 GAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 660  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720  
Db 661 ATAGAACTAGCCCAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 720  
Qy 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAA 780  
Db 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTTAA 780  
Qy 781 AAAAAATTTAAATTTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
Db 781 AAAAAATTTAAATTTTAAAGAAATTTTATTAACAAGTTGATTAAGATCTATTAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAATTA 873

## RESULT 9

US-10-303-162-13  
; Sequence 13, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-162-13

Query Match 99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCGTGAAATGGACCAAGTTTAAAGAAATGGATTATTCAGG 60
Db 1 ATGAAAAAGTTATTATTCGTGAAATGGACCAAGTTTAAAGAAATGGATTATTCAGG 60

QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTCGAAGATAATACTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTCGAAGATAATACTAT 120

QY 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATTAATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATTAATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAA 360

QY 361 TTTTCAGGAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTGCGAGTAGCC 420
Db 361 TTTTCAGGAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTGCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCCTCGCATATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600
Db 541 GATCCTCGCATATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATCTTACAAATAAATATATGCTTATGCTTAAATAGTCTTTTACCAATTTT 660
Db 601 GAAAAAATCTTACAAATAAATATATGCTTATGCTTAAATAGTCTTTTACCAATTTT 660

QY 661 ATAGAACTAGGCCCAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720

Db 661 ATAGAACTAGGCCCAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720
QY 721 AAAGATATACATCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Db 721 AAAGATATACATCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
QY 781 AAAAAATATAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATATAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGTATATAAAGCATTTATTTCAAAGGAAAA 873
Db 841 CCTAGTGTATATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 10
US-10-303-134-13
; Sequence 13, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-134-13

Query Match 99.3%; Score 869.8; DB 3; Length 873;
Best Local Similarity 99.8%; Pred. No. 8.9e-159;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCGTGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
Db 1 ATGAAAAAGTTATTATTCGTGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60

QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTCGAAGATAATACTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTCGAAGATAATACTAT 120

QY 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAGACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATTAATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATTAATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAA 360

QY 361 TTTTCAGGAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTGCGAGTAGCC 420
Db 361 TTTTCAGGAATTTATTTCAATCAAGAAATTAATCTCAGGGGTCTATATGTGCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCCTCGCATATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600
Db 541 GATCCTCGCATATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATCTTACAAATAAATATATGCTTATGCTTAAATAGTCTTTTACCAATTTT 660
Db 601 GAAAAAATCTTACAAATAAATATATGCTTATGCTTAAATAGTCTTTTACCAATTTT 660

QY 661 ATAGAACTAGGCCCAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720



Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360  
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATGCTCGCACTATATCGGACATAGTAAATAATACAGATATATAAGCTTTTAGAAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGACATAGTAAATAATACAGATATATAAGCTTTTAGAAATTTCTA 600  
Qy 601 GAAAAAATTTACAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACACT 720  
Db 661 ATAGAACTAGCCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACACT 720  
Qy 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAATAATTAATTTT 780  
Db 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAATAATTAATTTT 780  
Qy 781 AAAAAATATAAATTTAAAGAAATTTTATTTACAAAGTTGATATAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTTAAAGAAATTTTATTTACAAAGTTGATATAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

## RESULT 11

US-10-303-118-13  
; Sequence 13, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-118-13

Query Match

99.3%; Score 869.8; DB 3; Length 873;

Best Local Similarity 99.8%; Pred. No. 8.9e-159;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTTATTTATGCTGAAATGGACCAAGTTTAAAGAAAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTTATTTATGCTGAAATGGACCAAGTTTAAAGAAAAATTTGATTATTCAAGG 60  
Qy 61 CTACCAAAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAAATACAT 120  
Db 61 CTACCAAAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAAATACAT 120  
Qy 121 CTTGGTAAAAATGCAAGAGAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACATAC 180  
Db 121 CTTGGTAAAAATGCAAGAGAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACATAC 180  
Qy 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTGTTCTAAT 240  
Db 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATATTTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATATTTTTCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATGCTCGCACTATATCGGACATAGTAAATAATACAGATATAAGCTTTTAGAAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGACATAGTAAATAATACAGATATAAGCTTTTAGAAATTTCTA 600  
Qy 601 GAAAAAATTTACAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACACT 720  
Db 661 ATAGAACTAGCCGCAAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACACT 720  
Qy 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAATAATTAATTTT 780  
Db 721 AAAGATATACATACACTCTTAGTCAGGCTTATGGAATTTTCAAAATAATTAATTTT 780  
Qy 781 AAAAAATATAAATTTAAAGAAATTTTATTTACAAAGTTGATATAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATATAAATTTAAAGAAATTTTATTTACAAAGTTGATATAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

## RESULT 12

US-10-303-128-13  
; Sequence 13, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-128-13

Query Match
Best Local Similarity 99.3%; Score 869.8; DB 3; Length 873;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAATGATTTGATGATTAATGATGATCAATTTTAAAGAAATGATTATTCAAT 120
DB 61 CTACCAATGATTTGATGATTAATGATGATCAATTTTAAAGAAATGATTATTCAAT 120

QY 121 CTTGCTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCATACTAC 180
DB 121 CTTGCTAAAAATGCAAGCAGTGTTTACACCCCTAAATTTCTTCTTGAGCATACTAC 180

QY 181 ACTTTAAAAATTTAATCAAAATCAAGAAATGAGCCGAACTAATATGTTCTTAAT 240
DB 181 ACTTTAAAAATTTAATCAAAATCAAGAAATGAGCCGAACTAATATGTTCTTAAT 240

QY 241 TACACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT 300
DB 241 TACACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 360

QY 361 TTTACGAAATTTATTTCATCAAGAAATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 420
DB 361 TTTACGAAATTTATTTCATCAAGAAATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 480

QY 481 TCTTATGCTTTTGATACAAAGAAATCTTTTAAACAACTAGCCCTGATTTTAAAT 540
DB 481 TCTTATGCTTTTGATACAAAGAAATCTTTTAAACAACTAGCCCTGATTTTAAAT 540

QY 541 GATGCTCGCACTATATCGCATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATGCTCGCACTATATCGCATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTACAAATAAACTATTGCTTATGCTTATGCTTATGCTTTTACCAATTT 660
DB 601 GAAAAAATTTACAAATAAACTATTGCTTATGCTTATGCTTATGCTTTTACCAATTT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720

661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720
721 AAGATATATCTATACCTCTAGTGAGGCTTTATGAAAAATTTTCAAAAAATATTATTTT 780
721 AAGATATATCTATACCTCTAGTGAGGCTTTATGAAAAATTTTCAAAAAATATTATTTT 780
781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAGATT 840
781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAGATT 840
841 CCTAGTGATATAAAGCAATTTTCAAAAGGAAAA 873
841 CCTAGTGATATAAAGCAATTTTCAAAAGGAAAA 873

RESULT 13
US-10-735-419-13
; Sequence 13, Application US/10735419.
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-735-419-13

Query Match
Best Local Similarity 99.3%; Score 869.8; DB 5; Length 873;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAATGATTTGATGATTAATGATGATCAATTTTAAAGAAATGATTATTCAAT 120
DB 61 CTACCAATGATTTGATGATTAATGATGATCAATTTTAAAGAAATGATTATTCAAT 120

QY 121 CTTGCTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATAC 180
DB 121 CTTGCTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATAC 180

QY 181 ACTTTAAAAATTTAATCAAAATCAAGAAATGAGCCGAACTAATATGTTCTTAAT 240
DB 181 ACTTTAAAAATTTAATCAAAATCAAGAAATGAGCCGAACTAATATGTTCTTAAT 240

QY 241 TACACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT 300
DB 241 TACACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 360

QY 361 TTTACGAAATTTATTTCATCAAGAAATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 420
DB 361 TTTACGAAATTTATTTCATCAAGAAATTTTAAACAACTTAAAGAAATTTAAATGCTTATTAAA 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 480

QY 481 TCTTATGCTTTTGATACAAAGAAATCTTTTAAACAACTAGCCCTGATTTTAAAT 540
DB 481 TCTTATGCTTTTGATACAAAGAAATCTTTTAAACAACTAGCCCTGATTTTAAAT 540

QY 541 GATGCTCGCACTATATCGCATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATGCTCGCACTATATCGCATAGTAAATATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTACAAATAAACTATTGCTTATGCTTATGCTTATGCTTTTACCAATTT 660
DB 601 GAAAAAATTTACAAATAAACTATTGCTTATGCTTATGCTTATGCTTTTACCAATTT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATCAAGAAAAAATAACTACACT 720

Db 301 GATGCTCAATGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTTAA 360  
Qy 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
Qy 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAATAAACTATATGCTTTATGCTTATGCTTAAATTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAATAAACTATATGCTTTATGCTTAAATTTTAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 720  
Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 720  
Qy 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
Qy 781 AAAAAATTAATAAATAAAATTTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATTAATAAATAAAATTTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

## RESULT 14

US-09-816-028A-11  
; Sequence 11, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-09-816-028A-11

Query Match 98.9%; Score 866.6; DB 3; Length 873;  
Best Local Similarity 99.5%; Pred. No. 3.7e-158;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATCAAAAAAGTTTATTTATGCTGAAATGGACCAAGTTTAAAGAAATTTGATTTATTCAGG 60  
Db 1 ATCAAAAAAGTTTATTTATGCTGAAATGGACCAAGTTTAAAGAAATTTGATTTATTCAGG 60  
Qy 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGGAGCAATACTAC 180  
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTTGGAGCAATACTAC 180  
Qy 181 ACTTTAAAAATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATTTATGTGTCTAAT 240  
Db 181 ACTTTAAAAATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATTTATGTGTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Qy 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
Db 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCGCATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600  
Qy 601 GAAAAAATTTACAAATAAACTATATGCTTTATGCTTAAATTTTATGCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAATTTACAAATAAACTATATGCTTTATGCTTAAATTTTATGCTTTTAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 720  
Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 720  
Qy 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
Qy 781 AAAAAATTAATAAATAAAATTTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATTAATAAATAAAATTTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

## RESULT 15

US-10-303-162-11  
; Sequence 11, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162

; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-162-11

Query Match 98.9%; Score 866.6; DB 3; Length 873;  
Best Local Similarity 99.5%; Pred. No. 3.7e-158;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTC AAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTC AAGG 60  
QY 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGGAAGTAATACTAT 120  
Db 61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGGAAGTAATACTAT 120  
QY 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGTTTCTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAACTAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAACTAATATGTTCTTAAT 240  
QY 241 TACAACCGAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
Db 241 TACAACCGAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360  
QY 361 TTTTCAGAAAATTTAATCAAGAAATTTACCTCAGGGGTCTATATGTTGCAGTAGCC 420  
Db 361 TTTTCAGAAAATTTAATCAAGAAATTTACCTCAGGGGTCTATATGTTGCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAATAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAATAAT 540  
QY 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCGACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAAGCTTACAAATATAATGCTTATGCTTATGCTTAAATGCTTTTAGCAATTTT 660  
Db 601 GAAAAAAGCTTACAAATATAATGCTTATGCTTATGCTTAAATGCTTTTAGCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATAAAGAAAAAATACTACCT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATAAAGAAAAAATACTACCT 720  
QY 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780

Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAAATAAAAAATTTAAAGAAAAATGTTTATTACAAAGTTTGATAAAAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAAAAATTTAAAGAAAAATGTTTATTACAAAGTTTGATAAAAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATATAAGCATTTATTTCAAGGAAAA 873  
Db 841 CCTAGTGATATATAAGCATTTATTTCAAGGAAAA 873

Search completed: July 19, 2006, 15:09:23  
Job time : 138.645 secs

GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:30:11 ; Search time 137.396 Seconds  
(without alignments)  
1959.153 Million cell updates/sec

Title: US-10-734-719-12  
Perfect score: 1563  
Sequence: 1 MKKVIIAGNGSLKEIDYSR.....KLIKDLRLPSDIKHYFKGK 291

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1563	100.0	291	2	Q9F0M9_CAMJE	Q9F0M9 campylobact
2	1557	99.6	291	2	Q9L9Q5_CAMJE	Q9L9Q5 campylobact
3	1553	99.4	291	2	Q93D05_CAMJE	Q93D05 campylobact
4	1531	98.0	291	2	Q93M00_CAMJE	Q93M00 campylobact
5	1524	97.5	291	2	Q9LAK3_CAMJE	Q9LAK3 campylobact
6	1515	96.9	284	2	Q50F20_CAMJE	Q50F20 campylobact
7	1510	96.6	291	2	Q938X6_CAMJE	Q938X6 campylobact
8	1495	95.6	291	2	Q93C25_CAMJE	Q93C25 campylobact
9	1454.5	93.1	292	2	Q5DT12_CAMJE	Q5DT12 campylobact
10	1191.5	76.2	263	2	Q2TKA8_CAMJE	Q2TKA8 campylobact
11	804	51.4	294	2	Q7BP25_CAMJE	Q7BP25 campylobact
12	749	47.9	430	2	Q9RGFI_CAMJE	Q9RGFI campylobact
13	746.5	47.8	430	2	Q32VR8_CAMJE	Q32VR8 campylobact
14	743.5	47.6	326	2	Q4QM36_HAEI8	Q4QM36 haemophilus
15	729.5	46.7	320	2	Q4CNT8_HAEI8	Q4CNT8 haemophilus
16	698.5	44.7	303	2	Q5CLP3_PASMU	Q5CLP3 pasteurella
17	485.5	31.1	231	1	Y352_HAEIN	P24324 haemophilus
18	169	10.8	44	2	Q4HG71_CAMCO	Q4HG71 campylobact
19	151.5	9.7	1014	2	Q8RIY7_CAMJE	Q8RIY7 campylobact
20	151.5	9.7	1014	2	Q8RN21_CAMJE	Q8RN21 campylobact
21	150.5	9.6	1176	2	Q51CW9_ENTHI	Q51CW9 entamoeba h
22	143.5	9.2	673	2	Q4HEM3_CAMCO	Q4HEM3 campylobact
23	136	8.7	271	2	Q8NR31_SHEFR	Q8NR31 shewarella
24	131.5	8.4	2755	2	Q8IHP8_PLAF7	Q8IHP8 plasmidium
25	129.5	8.3	615	2	Q4XLN2_PLACH	Q4XLN2 plasmidium
26	129	8.3	849	2	Q8ISV7_PLAF7	Q8ISV7 plasmidium
27	129	8.3	1639	2	Q4YVD8_PLABE	Q4YVD8 plasmidium
28	128	8.2	819	2	Q6LFS2_PLAF7	Q6LFS2 plasmidium
29	128	8.2	2772	2	Q7RG21_PLAYO	Q7RG21 plasmidium
30	127.5	8.2	560	2	Q8IJN4_PLAF7	Q8IJN4 plasmidium
31	127.5	8.2	983	2	Q7RAH2_PLAYO	Q7RAH2 plasmidium

32	127	8.1	842	2	Q8ISV4_PLAF7	Q8ISV4 plasmidium
33	126	8.1	567	2	Q4XIM5_PLACH	Q4XIM5 plasmidium
34	125	8.0	713	2	Q6BT89_DEBHA	Q6BT89 debaryomyce
35	124	7.9	613	2	Q7V0X7_PROMP	Q7V0X7 prochloroco
36	124	7.9	1453	2	Q9G9H3_SCHCO	Q9G9H3 schizophyll
37	123.5	7.9	601	2	Q891W3_CLOTE	Q891W3 clostridium
38	123.5	7.9	803	2	Q4YST0_PLABE	Q4YST0 plasmidium
39	123	7.9	392	2	Q8FCQ6_ECOL6	Q8FCQ6 escherichia
40	123	7.9	417	2	Q9ALTO_CAMJE	Q9ALTO campylobact
41	123	7.9	475	2	Q4Z4E7_PLABE	Q4Z4E7 plasmidium
42	123	7.9	496	2	Q5CS63_CRYPV	Q5CS63 cryptospori
43	123	7.9	830	2	Q7RPT9_PLAYO	Q7RPT9 plasmidium
44	123	7.9	885	2	Q8IEM9_PLAF7	Q8IEM9 plasmidium
45	123	7.9	970	2	Q7RLC5_PLAYO	Q7RLC5 plasmidium

ALIGNMENTS

RESULT 1  
Q9F0M9 CAMJE  
ID Q9F0M9 CAMJE PRELIMINARY; PRT; 291 AA.  
AC Q9F0M9  
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.  
DT 01-MAR-2001, sequence version 1.  
DT 07-FEB-2006, entry version 12.  
DE Alpha-2,3-sialyltransferase.  
GN Name=cst-II;  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Campylobacteriaceae; Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=ATCC 43432;  
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NoDerivs License  
CC -----  
DR EMBL; AF215659; AAG43979.1; -; Genomic\_DNA.  
DR SMR; Q9F0M9; 1-258.  
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.  
DR InterPro; IPR009251; CST-I.  
DR Pfam; PF06002; CST-I; 1.  
KW Glycosyltransferase; Transferase.  
SQ SEQUENCE 291 AA; 34554 MW; B41B59ACD7280F8 CRC64;

Query Match 100.0%; Score 1563; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 4.8e-105;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKKVVIAGNGSLKEIDYSRLPNDVFRNCNQYFEDKYLGKKCKAVFVTPGFFQYY	60
Db	1	MKKVVIAGNGSLKEIDYSRLPNDVFRNCNQYFEDKYLGKKCKAVFVTPGFFQYY	60
QY	61	TLKHLIQOEYETELIMCSYNOAHLENENFVKTFFDYFDDAHLGYDFFKOLKEFNAYFK	120
Db	61	TLKHLIQOEYETELIMCSYNOAHLENENFVKTFFDYFDDAHLGYDFFKOLKEFNAYFK	120
QY	121	FHEIFVNRITSGVVCVAIAALGYKEIYLSGIDFQNGSSYAFDTTKOENLLKLAPFPKN	180
Db	121	FHEIFVNRITSGVVCVAIAALGYKEIYLSGIDFQNGSSYAFDTTKOENLLKLAPFPKN	180
QY	181	DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIEKKNYT	240
Db	181	DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIEKKNYT	240
QY	241	KDILIPSSSAYGKFSKNINFKIKIKENYVYKLIKOLLRLPSDIKHYFKGK	291
Db	241	KDILIPSSSAYGKFSKNINFKIKIKENYVYKLIKOLLRLPSDIKHYFKGK	291



```
Db 121 FHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFK 180
Qy 181 DRSHVIGHSKNTDIIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLNSFIIOEKNNYT 240
Db 181 DRSHVIGHSKNTDIIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLNSFIIOEKNNYT 240
Qy 241 KDILIPSSSEAYGKFSKNTNFKKIKENIVVYKLIKDLARLPSDIKHYPFGK 291
Db 241 KDILIPSSSEAYGKFSKNTNFKKIKENIVVYKLIKDLARLPSDIKHYPFGK 291

RESULT 4
Q93MQ0_CAMJE PRELIMINARY; PRT; 291 AA.
ID Q93MQ0;
AC Q93MQ0;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Alpha-2,3-/alpha-2,8-sialyltransferase.
GN Name=csII;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43438;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF400048; AAK91725.1; -; Genomic_DNA.
DR SRR; 993M00; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-1.
DR Pfam; PF06002; CST-1; 1.
DR KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34573 MW; BD1EDEF6F521E6 CRC64;

Query Match 98.0%; Score 1531; DB 2; Length 291;
Best Local Similarity 97.9%; Pred. No. 1e-102;
Matches 285; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLRPNDFVPRCNOQFYFEDKYLGKCKAVFYTPGPFQYV 60
Db 1 MKKVIIAGNGPSLKEIDYSLRPNDFVPRCNOQFYFEDKYLGKCKAVFYTPGPFQYV 60
Qy 61 TLKHLIQNOEYETELIMCSNTYNOAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSNTYNOAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120
Qy 121 FHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFK 180
Db 121 FHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFK 180
Qy 181 DRSHVIGHSKNTDIIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLNSFIIOEKNNYT 240
Db 181 DRSHVIGHSKNTDIIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLNSFIIOEKNNYT 240
Qy 241 KDILIPSSSEAYGKFSKNTNFKKIKENIVVYKLIKDLARLPSDIKHYPFGK 291
Db 241 KDILIPSSSEAYGKFSKNTNFKKIKENIVVYKLIKDLARLPSDIKHYPFGK 291

RESULT 5
Q9LAK3_CAMJE PRELIMINARY; PRT; 291 AA.
ID Q9LAK3;
AC Q9LAK3;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Alpha-2,3-/8-sialyltransferase (Alpha-2,3-/2,8-sialyltransferase).
GN Name=csII;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=OH4384;
RC MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brissou J.-R., Karwaski M.-F., Michniewicz J.,
RA Cunningham A.-M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT MHz (1)H and (13)C NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 700297;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=CF 93-6;
RC Koga M., Gilbert M., Li J., Koike S., Takahashi M., Furukawa K.,
RA Hirata K., Yuki N.;
RT "Antecedent infections in Fisher Syndrome: a common pathogenesis of
RT molecular mimicry.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF130984; AAF31771.1; -; Genomic_DNA.
DR EMBL; AF218647; AAL3462.1; -; Genomic_DNA.
DR EMBL; AY644679; AAV52039.1; -; Genomic_DNA.
DR PDB; 1RO7; X-ray; A/B/C/D=1-259.
DR PDB; 1RO8; X-ray; A/B=1-259.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-1.
DR Pfam; PF06002; CST-1; 1.
DR KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34545 MW; AB0741D5FCEDAE6A CRC64;

Query Match 97.5%; Score 1524; DB 2; Length 291;
Best Local Similarity 97.3%; Pred. No. 3.2e-102;
Matches 283; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLRPNDFVPRCNOQFYFEDKYLGKCKAVFYTPGPFQYV 60
Db 1 MKKVIIAGNGPSLKEIDYSLRPNDFVPRCNOQFYFEDKYLGKCKAVFYTPGPFQYV 60
Qy 61 TLKHLIQNOEYETELIMCSNTYNOAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120
Db 61 TLKHLIQNOEYETELIMCSNTYNOAHLENENFVKTFYDYPDAHLGYDFFKQKEFNAYFK 120
Qy 121 FHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFK 180
Db 121 FHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPFK 180
Qy 181 DRSHVIGHSKNTDIIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLNSFIIOEKNNYT 240
Db 181 DRSHVIGHSKNTDIIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLNSFIIOEKNNYT 240
Qy 241 KDILIPSSSEAYGKFSKNTNFKKIKENIVVYKLIKDLARLPSDIKHYPFGK 291
Db 241 KDILIPSSSEAYGKFSKNTNFKKIKENIVVYKLIKDLARLPSDIKHYPFGK 291
```

```

RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC
DR EMBL; AY044868; AAK96001.1; -; Genomic_DNA.
DR SMR; Q938X6; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34577 MW; 7BAE5P6021A56F08 CRC64;

Query Match 96.6%; Score 1510; DB 2; Length 291;
Best Local Similarity 95.5%; Pred. No. 3.3e-101;
Matches 278; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNQFYEDKYILGKCKKAVFYTPGFFPEQYY 60
DB 1 MKKVIIAGNGPSLKEIDYSRLPNDPFRNQFYEDKYILGKCKKAVFYTPGFFPEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
DB 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 FHEIFYNQRTISGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
DB 121 FHEIFYNQRTISGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHYPKKG 291
DB 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHYPKKG 291

RESULT 8
Q93CZ5 CAMJE PRELIMINARY; PRT; 291 AA.
AC Q93CZ5;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Bifunctional alpha-2,3/-2,8-sialyltransferase.
GN Name=est-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RT ganglioside mimics in the core oligosaccharide.";
RL J. Biol. Chem. 277:327-337(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC
DR EMBL; AF401529; AAL06004.1; -; Genomic_DNA.
DR SMR; Q93CZ5; 1-258.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.

```



```

DR InterPro: IPR009251; CST-I.
DR Pfam: PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34505 MW; 50162DEE8D81558F CRC64;

Query Match
Best Local Similarity 95.6%; Score 1495; DB 2; Length 291;
Matches 275; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTPGFFPEQY 60
Db 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTPGFFPEQY 60

QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120

QY 121 PHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 PHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

QY 291 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
Db 291 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 9
QSDT12 CAMJE
ID QSDT12 CAMJE PRELIMINARY; PRT; 292 AA.
AC QSDT12;
DT 29-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 29-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Putative alpha-2,3/-2,8 sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MF6;
RA Gilbert M., Karwaski M.-F., Godschalk P.C.R., Brochu D., Endtz H.P.,
RA Cunningham A.;
RT "Sequencing of the lipooligosaccharide biosynthesis locus of
RT Campylobacter jejuni MF6."
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AY422196; AAR82857.1; -; Genomic DNA.
DR SMR; QSDT12; 1-259.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro: IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 292 AA; 34658 MW; 1A6F7B6FF7B0A2F CRC64;

Query Match
Best Local Similarity 93.1%; Score 1454.5; DB 2; Length 292;
Matches 269; Conservative 12; Mismatches 10; Indels 1; Gaps 1;

QY 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTPGFFPEQY 60
Db 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTPGFFPEQY 60

QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120

```

```

QY 121 PHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 PHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 239
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240

QY 240 TKDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 291
Db 241 TKDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFGK 292

RESULT 10
Q2TKA8 CAMJE
ID Q2TKA8 CAMJE PRELIMINARY; PRT; 263 AA.
AC Q2TKA8;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Putative alpha-2,3/8-sialyltransferase.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GB5;
RA Gilbert M., Brochu D., Karwaski M.-F.;
RA "Sequencing of the Campylobacter jejuni GB5 LOS locus." ;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
EMBL; AY854153; AAX45339.1; -; Genomic DNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 263 AA; 31295 MW; 21E028981B9823E0 CRC64;

Query Match
Best Local Similarity 76.2%; Score 1191.5; DB 2; Length 263;
Matches 224; Conservative 13; Mismatches 17; Indels 13; Gaps 2;

QY 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTPGFFPEQY 60
Db 1 MKKVIIAGNPSLKEIDYSLRPNDFVRCNQFYFEDKYLGKCKKAVFYTPGFFPEQY 60

QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFKQKKEFNAYFK 120

QY 121 PHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 PHEIYNQRIITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240

QY 241 KDI-----LIPSSAYGKFSKNINFKKI 263
Db 241 LKIIYLLVRLME-----NFQKI 258

RESULT 11
Q7BP25 CAMJE
ID Q7BP25 CAMJE PRELIMINARY; PRT; 294 AA.
AC Q7BP25; Q9PNF4;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.

```

[illegible]

```

RESULT 12
Q9RGF1_CAMUJE
ID Q9RGF1_CAMJE PRELIMINARY; PRT; 430 AA.
AC Q9RGF1;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Alpha-2,3-sialyltransferase.
GN Name=cst-I;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OH4384;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brisson J.-R., Karwaski M.-E., Michniewicz J.,
RA Cunningham A.-M., Wu Y., Young N.-M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RL MHz (1)H and (13)C NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
RJ [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OH4384;
RA Gilbert M., Wakarchuk W.W.;
RT "Lipopolysaccharide .alpha.-2,3 sialyltransferase of campylobacter.";
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
CC EMBL; AF130466; AAF13495.1; -; Genomic DNA.
CC GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
CC InterPro; IPR009251; CST-I.
CC Pfam; PF06002; CST-I; 1.
CC Glycosyltransferase; Transferase.
SQ SEQUENCE 430 AA; 50396 MW; 1D03B6797169425C CRC64;

```

Query Match            47.9%; Score 749; DB 2; Length 430;  
Best Local Similarity 50.9%; pred. No. 6.3e-46;  
Matches 148; Conservative 45; Mismatches 80; Indels 18; Gaps 5

QY     1 MKKVIAGNPSLKEIDYSRLPNDFVRCNQFYEDKYILGKKCAVFYTGFFFEQYY 60  
     : :::::::::::::::::::::  
     : :::::::::::::::::::::

16	MQNLLIAGNGFSUKNNINIKKLPFLBIDYDFKUNQFYFEDUKYILGKKLKAFFNFNPGVDFLQQH	15
61	TLKHLIQLONQYETELIMCSNYNOAHLENENFVKTFYDFYDPAHLGSDYFFKQLKEFNAYPK	120
76	TAKQILKNKEYEIKNIFCSTFNLPFIESNDPLHQFYFNFPFPAKLGVEYIENLKEFYVYIK	135
121	FHEIYENORITSGVYWCVAIAIAGYKYLIGSDIFYQNGSYPADTKOENLLKLP--D	177
136	YNEIYFNKRITSGVYMCIAIAGYKTYILCGDIFYEGDVTIYPPEAMSTNKTITPFGIKD	195
178	FKNDRSHYIGHSKNTDIKALEFEKTYKIKLYCLCPNSLLANFIELAPNLNSNFITQEK-	236
196	FKPSNC---HSKEYDIEALKLAKSIYKVNIVYALCDDSLANHPFLSININNFNTELENKH	251

[illegible]

QY 237 NNYTKDILIPSSSEAYGKESKN-----INFKKI-KIKENVYKLIKDL 277  
 DB 252 NNSINDILLTDNTPGVSPYKQKADNKNIMLNFYNILHSDKNILKFLNKEI 302

RESULT 13  
 Q32VR8 CAMJE PRELIMINARY; PRT; 430 AA.  
 AC Q32VR8;  
 DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.  
 DT 06-DEC-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 3.  
 DE Alpha-2,3-sialyltransferase.  
 GN Name=stt1;  
 OS Campylobacter jejuni subsp. jejuni.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=32022;  
 RN [1]

RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 43446;  
 RA Gilbert M., Brochu D., Karwaski M.-F.;  
 RT "Sequencing of the Cst-I locus of Campylobacter jejuni ATCC 43446.";  
 RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC EMBL; AY791515; AAX33819.1; -; Genomic DNA.  
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.  
 KW Glycosyltransferase; Transferase.  
 SQ SEQUENCE 430 AA; 50383 MW; 876F9E0583694240 CRC64;

Query Match 47.8%; Score 746.5; DB 2; Length 430;  
 Best Local Similarity 52.7%; Pred. No. 9.5e-46;  
 Matches 149; Conservative 40; Mismatches 85; Indels 9; Gaps 4;

QY 1 MKVIIAGNGPSLKEIDYSLPNDVDVFRGNQFVEDKYLGKCKAVFYTPGFFFEQYV 60  
 DB 16 MQNIIAGNGPSLKNINYLPRYDVFRCNQFVEDKYLGKCKAVFYTPGFFFEQYH 75  
 QY 61 TLKHLIQNQYETELIMCSNNOAHLENENFVKTFDYDPPDAHLGYDFKQKFNAYFK 120  
 DB 76 TAKQLIKNEYEIKNIFCSTNLPFIESNDLHQFYNFPDPAKLGYEVLENLKEFYAIK 135  
 QY 121 FHEIFYNQRTSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKAP 177  
 DB 136 YNEIFYNKRITSGVYMCVAIAIALGYKTYILCGIDFYEGDVIYPPFAMSTNKTIPP 195  
 QY 178 FKNDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFTELAPNLSNFIQEK 236  
 DB 196 FKPSNC-----HSKEYDIEALKLSYKVIYALCDDSIILANHPFLSININNFTLENKH 251  
 QY 237 NNYTKDILIPSSSEAYGKESKNINFKKIKIKENVYKLIK-IKDLL 278  
 DB 252 NNSINDILLTDNTPGVSPYKQKADNKNIMLNFYNILHSDKNILHSDKDTL 294

RESULT 14  
 Q4QM36 HAEI8 PRELIMINARY; PRT; 326 AA.  
 AC Q4QM36;  
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.  
 DT 19-JUL-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 5.  
 DE CMP-Neu5Ac--lipooligosaccharide alpha 2-3 sialyltransferase.  
 GN Name=lic3A2; OrderedLocusNames=NTH11034;  
 OS Haemophilus influenzae (strain 86-028NP);  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=281310;  
 RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;  
 RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,  
 RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,  
 RA Munson R.S. Jr.;  
 RT "Genomic sequence of an otitis media isolate of nontypeable  
 RT Haemophilus influenzae: comparative study with H. influenzae serotype  
 RT d, strain KW20.";  
 RL J. Bacteriol. 187:4627-4636(2005).  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC EMBL; CP000057; AAX87911.1; -; Genomic DNA.  
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.  
 KW Complete proteome; Glycosyltransferase; Transferase.  
 SQ SEQUENCE 326 AA; 38272 MW; 78FF283EB10C6F0E CRC64;

Query Match 47.6%; Score 743.5; DB 2; Length 326;  
 Best Local Similarity 50.7%; Pred. No. 1.1e-45;  
 Matches 149; Conservative 49; Mismatches 87; Indels 9; Gaps 6;

QY 2 KKVIIAGNGPSLKEIDYSELNDVDVFRGNQFVEDKYLGKCKAVFYTPGFFFEQYV 61  
 DB 32 KSVIIAGNGTSLSKIDYSLPNDVDVFRGNQFVEDKYLGKCKAVFYTPGFFFEQYV 91  
 QY 62 LKHLIQNQYETELIMCSNNOAHLENENFVKTFDYDPPDAHLGYDFKQKFNAYFK 120  
 DB 92 FMQLIKNEYEVADIILSSFLNLGSELK--TORLEKLIPQIDLGHSLKLAFAFDAQ 150  
 QY 121 FHEIFYNQRTSGVYMCVAIAIALGYKEIYLSGIDFYQ--NGSSYAFDTKQENLLKAPDFK 179  
 DB 151 YHELYENKRITSGVYMCVAIAIALGYKDYLTGIDFYQKGNPYAFHQTENIILKLPSPS 210  
 QY 180 NDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFTELAPNLS--NFITQEK 237  
 DB 211 QNKQSDIHSMEYDNLALYFLQKHYGVNICYSPSPCLCNFYPLSP--LNNPITFILEKK 269  
 QY 238 NYTKDILIPSSSEAYGKESKNINFKKIKIKENVYKLIKIDLLRPSDKHYFKCK 291  
 DB 270 NYTDILIPPKFYKIG---IYKPRIYQNLIFRLINDILRLPNDIKHAKKAK 320

RESULT 15  
 Q4QNI8 HAEI8 PRELIMINARY; PRT; 320 AA.  
 AC Q4QNI8;  
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.  
 DT 19-JUL-2005, sequence version 1.  
 DT 07-FEB-2006, entry version 5.  
 DE CMP-Neu5Ac--lipooligosaccharide alpha 2-3 sialyltransferase.  
 GN Name=lic3A; OrderedLocusNames=NTH10472;  
 OS Haemophilus influenzae (strain 86-028NP).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=281310;  
 RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
 RX PubMed=15968074; DOI=10.1128/JB.187.13.4627-4636.2005;  
 RA Harrison A., Dyer D.W., Gillaspay A., Ray W.C., Mungur R., Carson M.B.,  
 RA Zhong H., Gipson J., Gipson M., Johnson L.S., Lewis L., Bakaletz L.O.,  
 RA Munson R.S. Jr.;  
 RT "Genomic sequence of an otitis media isolate of nontypeable  
 RT Haemophilus influenzae: comparative study with H. influenzae serotype  
 RT d, strain KW20.";  
 RL J. Bacteriol. 187:4627-4636(2005).  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NoDerivs License  
 CC EMBL; CP000057; AAX87409.1; -; Genomic DNA.  
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.  
 KW Complete proteome; Glycosyltransferase; Transferase.



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:47:11 ; Search time 27.928 Seconds  
(without alignments)  
912.038 Million cell updates/sec

Title: US-10-734-719-12

Perfect score: 1563

Sequence: 1 MKKVIAGNPSLKEIDYSR.....KLIKDLLRLPSDKIKYFGK 291

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/5\_COMB.pep.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/6\_COMB.pep.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/7\_COMB.pep.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/8\_COMB.pep.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/9\_COMB.pep.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/RE\_COMB.pep.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1563	100.0	291	2	US-09-816-028A-12
2	1563	100.0	291	2	US-10-303-162-12
3	1563	100.0	291	2	US-10-303-134-12
4	1563	100.0	291	2	US-10-303-118-12
5	1563	100.0	291	2	US-10-303-128-12
6	1563	100.0	291	3	US-10-735-419-12
7	1557	99.6	291	2	US-09-495-406-9
8	1557	99.6	291	2	US-09-816-028A-9
9	1557	99.6	291	2	US-10-303-162-9
10	1557	99.6	291	2	US-10-303-134-9
11	1557	99.6	291	2	US-10-303-118-9
12	1557	99.6	291	2	US-10-303-128-9
13	1557	99.6	291	3	US-10-735-419-9
14	1553	99.4	291	2	US-09-816-028A-14
15	1553	99.4	291	2	US-10-303-162-14
16	1553	99.4	291	2	US-10-303-134-14
17	1553	99.4	291	2	US-10-303-118-14
18	1553	99.4	291	2	US-10-303-128-14
19	1534	98.1	291	3	US-10-735-419-14
20	1534	98.1	291	2	US-09-495-406-5
21	1534	98.1	291	2	US-09-816-028A-5
22	1534	98.1	291	2	US-10-303-162-5
23	1534	98.1	291	2	US-10-303-134-5
24	1534	98.1	291	2	US-10-303-118-5
25	1534	98.1	291	2	US-10-303-128-5
26	1534	98.1	291	3	US-10-735-419-5

27	1524	97.5	291	2	US-09-495-406-3	Sequence 3, Appli
28	1524	97.5	291	2	US-09-816-028A-3	Sequence 3, Appli
29	1524	97.5	291	2	US-10-303-162-3	Sequence 3, Appli
30	1524	97.5	291	2	US-10-303-134-3	Sequence 3, Appli
31	1524	97.5	291	2	US-10-303-118-3	Sequence 3, Appli
32	1524	97.5	291	2	US-10-303-128-3	Sequence 3, Appli
33	1524	97.5	291	3	US-10-735-419-3	Sequence 3, Appli
34	1513	96.8	291	2	US-09-495-406-7	Sequence 7, Appli
35	1513	96.8	291	2	US-09-816-028A-7	Sequence 7, Appli
36	1513	96.8	291	2	US-10-303-162-7	Sequence 7, Appli
37	1513	96.8	291	2	US-10-303-134-7	Sequence 7, Appli
38	1513	96.8	291	2	US-10-303-118-7	Sequence 7, Appli
39	1513	96.8	291	2	US-10-303-128-7	Sequence 7, Appli
40	1513	96.8	291	3	US-10-735-419-7	Sequence 7, Appli
41	804	51.4	294	2	US-09-495-406-10	Sequence 10, Appl
42	804	51.4	294	2	US-09-816-028A-10	Sequence 10, Appl
43	804	51.4	294	2	US-10-303-162-10	Sequence 10, Appl
44	804	51.4	294	2	US-10-303-134-10	Sequence 10, Appl
45	804	51.4	294	2	US-10-303-118-10	Sequence 10, Appl

## ALIGNMENTS

### RESULT 1

US-09-816-028A-12  
; Sequence 12, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/1118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-09-816-028A-12

Query Match	100.0%;	Score	1563;	DB	2;	Length	291;
Best Local Similarity	100.0%;	Pred. No.	3.3e-155;				
Matches	291;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MKKVIAGNPSLKEIDYSRLPNDVFCNQFVEDKYLGKKCAVFTYTPGFFEQYY	60				
Db	1	MKKVIAGNPSLKEIDYSRLPNDVFCNQFVEDKYLGKKCAVFTYTPGFFEQYY	60				
QY	61	TLKHILQOEYETELIMCSYNOAHLENENFVKTFYDFPDHLYGDFDFKQKEFNAYFK	120				
Db	61	TLKHILQOEYETELIMCSYNOAHLENENFVKTFYDFPDHLYGDFDFKQKEFNAYFK	120				
QY	121	PHIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYONGSSYAPDTTQENLLKLAPDPKN	180				
Db	121	PHIYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYONGSSYAPDTTQENLLKLAPDPKN	180				
QY	181	DRSHYGHSKNTDIIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240				
Db	181	DRSHYGHSKNTDIIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240				
QY	241	KDILIPSSAYGKFSKNINFKIKIKENVVYKLIKILLRLPSDIKHFKGK	291				



Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSLRPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSLRPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120  
DB 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 PHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 PHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

## RESULT 5

US-10-303-128-12  
; Sequence 12, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-128-12

Query Match 100.0%; Score 1563; DB 2; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.3e-155;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSLRPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSLRPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120  
DB 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 PHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 PHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

RESULT 6  
US-10-735-419-12  
; Sequence 12, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-735-419-12

Query Match 100.0%; Score 1563; DB 3; Length 291;  
Best Local Similarity 100.0%; Pred. No. 3.3e-155;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSLRPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSLRPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60

QY 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120  
DB 61 TLKHLIQNOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

QY 121 PHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 PHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHFKGK 291

## RESULT 7

US-09-495-406-9  
; Sequence 9, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; CURRENT FILING DATE: 2000-01-31

```

; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRN
; ORGANISM: Campylobacter jejuni
; US-09-495,406-9

```

Query Match 99.6%; Score 1557; DB 2; Length 291;  
Best Local Similarity 99.7%; Pred. NO. 1.4e-154;  
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

61	QY	TLKHLIQNEYETELIMCSNYNQAHLENENFVKTYFYDPPDAHLGYDFFKQKLFENAYPK	120
61	Db	TLKHLIQNEYETELIMCSNYNQAHLENENFVKTYFYDPPDAHLGYDFFKQKLFENAYPK	120
121	QY	FHEIYFNORITSGVYMCVAIALGKETYLSGIDFYQNGSSYADPTKQENLLKLAADPKN	180
121	Db	FHEIYFNORITSGVYMCVAIALGKETYLSGIDFYQNGSSYADPTKQENLLKLAADPKN	180
181	QY	DRSHYIGHSKNTDICALFELEKTKYKLYCUCPSNLLANFTELPNLSNFTIOBKNVYT	240
181	Db	DRSHYIGHSKNTDICALFELEKTKYKLYCUCPSNLLANFTELPNLSNFTIOBKNVYT	240
241	QY	KDILIPSSSEAYGKPSKNINFKKIKIKGNVYKLIKOLLRLPSDINKHYFGKG	291
241	Db	KDILIPSSSEAYGKPSKNINFKKIKIKGNVYKLIKOLLRLPSDINKHYFGKG	291

```

RESULT 8
US-09-816-028A-9
; Sequence 9, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-09-816-028A-9

```

	Query Match	99.6%	Score 1557;	DB 2;	Length 291;
	Best Local Similarity	99.7%;	Pred. No. 1.4e-154;		
	Matches 290;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	MKKVVIAGNPSLKEIDYSLPLNDFVFCNQYFEDKYVLGKKCAKAVFTBGFPEQY	60		
Db	1	MKKVVIAGNPSLKEIDYSLPLNDFVFCNQYFEDKYVLGKKCAKAVFTBGFPEQY	60		
QY	61	TLKHLIQOEYETELIMCSNYNAHLENENFVKTIFYDYPDAHLGYDFFKQLKEFNAYFK	120		

61	TLKHLIQOEYETELIMCSNVNQAHLENENFWKTFYDFPDHILGVDFPKQLKFENAYFK	120
Db		
121	FHEIYFNQRTISGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKOENLLKLAPDFKN	180
Qy		
121	FHEIYFNQRTISGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKOENLLKLAPDFKN	180
Db		
121	FHEIYFNQRTISGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKOENLLKLAPDFKN	180
Qy		
181	DRSHYIGHSKNNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIIQEKNNYT	240
Db		
181	DRSHYIGHSKNNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPLNSNFIIQEKNNYT	240
Qy		
241	KDILIPSSBAYGFKSKNINFKKIKIKENVYKLIKDLLRLPSDIIKHYPGKG	291
Db		
241	KDILIPSSBAYGFKSKNINFKKIKIKENVYKLIKDLLRLPSDIIKHYPGKG	291

RESULT 9  
 US-10-303-162-9  
 ; Sequence 9, Application US/10303162  
 ; Patent No. 6723545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-00011US  
 ; CURRENT APPLICATION NUMBER: US/10/303,162  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 9  
 ; LENGTH: 291  
 ; TYPE: PRT  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 ; US-10-303-162-9

	Query Match	99.6%; Score 1557; DB 2; Length 291;
	Best Local Similarity	99.7%; Pred. No. 1.4e-154;
	Matches 290; Conservative	0; Mismatches 1; Indels 0; Gaps 0
Qy	1 MKKVIITAGNGPSLKEIDYSRLPNDFDVFRCNQFYFEDKYILGKKCAKAVYTPPGFFFEQYY	60
Dd		
1	1 MKKVIITAGNGPSLKEIDYSRLPNDFDVFRCNQFYFEDKYILGKKCAKAVYTPPNPFEEQYY	60
Qy	61 TLKHLIQNQVEYTELIMCSNQNQAHLNENPVKTIFYDPPDAHLYGDFFKQLKENAYPK	120
Dd	61 TLKHLIQNQVEYTELIMCSNQNQAHLNENPVKTIFYDPPDAHLYGDFFKQLKENAYPK	120
Qy	121 FHEIYNQRITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFPTKQENLLKLAPDFKN	180
Dd		
121	1 FHEIYNQRITSGVMCAVAIALGYKEIYLSGIDFYQNGSSYAFPTKQENLLKLAPDFKN	180
Qy	181 DRSHYIGHSNNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNVT	240
Dd	181 DRSHYIGHSNNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSNFIIQEKNNVT	240
Qy	241 KOILLPSSEAYGFSPKSNINFKKIKIKENVVYKLIKDLLRPLSDIKHYFGKG	291
Dd	241 KOILLPSSEAYGFSPKSNINFKKIKIKENVVYKLIKDLLRPLSDIKHYFGKG	291

RESULT 10  
US-10-303-134-9  
; Sequence 9, Application US/10303134



```

Query Match      99.6%; Score 1557; DB 2; Length 291;
Best Local Similarity 99.7%; Pred. No. 1.4e-154;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRLPNDFVFRNCNFYFEDKYILGKKCKAVFYTPGFEEQY 60
      |||||
Db 1 MKKVIAGNGPSLKEIDYSRLPNDFVFRNCNFYFEDKYILGKKCKAVFYTPNFEEQY 60
      |||||

QY 61 TLKHLIQNQEYETELIMCSNYNQAHLNENFVKTFYDFYFPAHLGYDFFKQLKEFNAYFK 120

```

Db 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDFPDHGLGYDFFKQKFNAYFK 120  
Qy 121 FHEIFYNQRITSGVYMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIFYNQRITSGVYMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Qy 181 DRSHYGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Qy 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291  
Db 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291

RESULT 13  
US-10-735-419-9  
; Sequence 9, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-735-419-9

Query Match 99.6%; Score 1557; DB 3; Length 291;  
Best Local Similarity 99.7%; Pred. No. 1.4e-154;  
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
Qy 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDFPDHGLGYDFFKQKFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDFPDHGLGYDFFKQKFNAYFK 120  
Qy 121 FHEIFYNQRITSGVYMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIFYNQRITSGVYMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Qy 181 DRSHYGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Qy 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291  
Db 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291

RESULT 14  
US-09-816-028A-14

; Sequence 14, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-09-816-028A-14

Query Match 99.4%; Score 1553; DB 2; Length 291;  
Best Local Similarity 99.3%; Pred. No. 3.7e-154;  
Matches 289; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSLPNDPFDVFCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60  
Qy 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDFPDHGLGYDFFKQKFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLNENFVKTFYDFPDHGLGYDFFKQKFNAYFK 120  
Qy 121 FHEIFYNQRITSGVYMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIFYNQRITSGVYMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Qy 181 DRSHYGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Qy 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291  
Db 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLRLPSDIKHVFKGK 291

RESULT 15  
US-10-162-14

; Sequence 14, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14

```
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-162-14

Query Match          99.4%; Score 1553; DB 2; Length 291;
Best Local Similarity 99.3%; Pred. NO. 3.7e-154;
Matches 289; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDVFRCNQFYFEDKYVLGKKCKAVFYTPGFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSLPNDVDVFRCNQFYFEDKYVLGKKCKAVFYTPGFFFEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120

QY 121 FHEIYFNQRTSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIYFNQRTSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDILLPSDIKHYPKKG 291
Db 241 KDILIPSEAYGKFSKNINFKKIKIKENYVYKLIKDILLPSDIKHYPKKG 291
```

Search completed: July 19, 2006, 13:50:05  
Job time : 27.928 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM protein - protein search, using sw model  
Run on: July 19, 2006, 14:13:03 ; Search time 95.0051 Seconds  
(without alignments)  
1418.823 Million cell updates/sec  
Title: US-10-734-719-12  
Perfect score: 1563  
Sequence: 1 MKKVIIAGNGSLKEIDYSR.....KLIKDLLRLPSDIKHVFKGK 291  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	291	3	US-09-816-028A-12
2	1563	100.0	291	4	US-10-303-161-12
3	1563	100.0	291	4	US-10-303-118-12
4	1563	100.0	291	4	US-10-303-128-12
5	1563	100.0	291	4	US-10-303-134-12
6	1563	100.0	291	4	US-10-303-162-12
7	1563	100.0	291	4	US-10-820-536-12
8	1563	100.0	291	4	US-10-845-408-12
9	1563	100.0	291	4	US-10-845-412-12
10	1563	100.0	291	5	US-10-846-219-12
11	1563	100.0	291	5	US-10-821-604-12
12	1563	100.0	291	5	US-10-847-983-12
13	1563	100.0	291	5	US-10-821-573-12
14	1563	100.0	291	5	US-10-850-807-12
15	1563	100.0	291	5	US-10-850-125-12
16	1563	100.0	291	5	US-10-830-825-12
17	1563	100.0	291	5	US-10-962-334-12
18	1563	100.0	291	5	US-10-830-997-12
19	1563	100.0	291	5	US-10-962-235-12
20	1563	100.0	291	5	US-10-961-882-12
21	1557	99.6	291	3	US-09-816-028A-9
22	1557	99.6	291	4	US-10-303-161-9
23	1557	99.6	291	4	US-10-303-118-9
24	1557	99.6	291	4	US-10-303-128-9
25	1557	99.6	291	4	US-10-303-134-9
26	1557	99.6	291	4	US-10-303-162-9
27	1557	99.6	291	4	US-10-820-536-9

ALIGNMENTS

RESULT 1

US-09-816-028A-12  
; Sequence 12, Application US/09816028A  
; Patent No. US20020042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Makarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-09-816-028A-12

Query Match	100.0%;	Score 1563;	DB 3;	Length 291;
Best Local Similarity	100.0%;	Pred. No. 2.1e-123;		
Matches 291;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKKVIIAGNGSLKEIDYSRLPNDVFRQNFYEDKYLYGKKCAVFTYTPGFFFEQY	60	
Db	1	MKKVIIAGNGSLKEIDYSRLPNDVFRQNFYEDKYLYGKKCAVFTYTPGFFFEQY	60	
Qy	61	TLKHLIQNEYETELIMCSNNOAHLENENFVKTFYDYFDDAHLGYDFFKQKEFNAYFK	120	
Db	61	TLKHLIQNEYETELIMCSNNOAHLENENFVKTFYDYFDDAHLGYDFFKQKEFNAYFK	120	
Qy	121	FHEIFNQBITSGVYMCVAIAIYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN	180	
Db	121	FHEIFNQBITSGVYMCVAIAIYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN	180	
Qy	181	DRSHYIGHSKNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240	
Db	181	DRSHYIGHSKNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT	240	
Qy	241	KDILIPSSAYGKFSKNINFKIKIKENVYKLIKILKLLRPLSDIKHYFKGK	291	

```
Db      241 KDILPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
;
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
;
; FEATURE:
;
; TITLE OF INVENTION: Campylobacter glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
;
; FEATURE:
;
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
; US-10-303-161-12

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYEDKYLGKCKAVFTPGFFFEQYY 60
Db      1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYEDKYLGKCKAVFTPGFFFEQYY 60
QY      61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTYDYPDAHLGYDFPKQKEFNAYFK 120
Db      61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTYDYPDAHLGYDFPKQKEFNAYFK 120
QY      121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db      121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY      181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY      241 KDILPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
Db      241 KDILPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 4
US-10-303-128-12
; Sequence 12, Application US/10303118
; Publication No. US20030157656A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
;
; FEATURE:
;
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
; US-10-303-128-12

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYEDKYLGKCKAVFTPGFFFEQYY 60
Db      1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYEDKYLGKCKAVFTPGFFFEQYY 60
QY      61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTYDYPDAHLGYDFPKQKEFNAYFK 120
Db      61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTYDYPDAHLGYDFPKQKEFNAYFK 120
QY      121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db      121 FHEIFYNQRITSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
QY      181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db      181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
QY      241 KDILPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291
Db      241 KDILPSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 3
US-10-303-118-12
; Sequence 12, Application US/10303118
; Publication No. US20030157655A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
```

QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYEDKYLGKCKKAVFTPGFFFEQY 60  
DB 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYEDKYLGKCKKAVFTPGFFFEQY 60  
QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120  
QY 121 FHEIFYNQRTITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
DB 121 FHEIFYNQRTITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291  
DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 5  
US-10-303-134-12  
; Sequence 12, Application US/10303134  
; Publication No. US20030157657A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-134-12

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.1e-123;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYEDKYLGKCKKAVFTPGFFFEQY 60  
DB 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYEDKYLGKCKKAVFTPGFFFEQY 60  
QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120  
QY 121 FHEIFYNQRTITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
DB 121 FHEIFYNQRTITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291  
RESULT 6  
US-10-303-162-12  
; Sequence 12, Application US/10303162  
; Publication No. US20030157659A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 291  
; TYPE: PRT  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-162-12

Query Match 100.0%; Score 1563; DB 4; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.1e-123;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYEDKYLGKCKKAVFTPGFFFEQY 60  
DB 1 MKKVIIAGNGPSLKEIDYSLPNDVDFRCNQFYEDKYLGKCKKAVFTPGFFFEQY 60  
QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120  
QY 121 FHEIFYNQRTITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
DB 121 FHEIFYNQRTITSGVVMCAVAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291  
DB 241 KDILIPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 7  
US-10-820-536-12  
; Sequence 12, Application US/10820536  
; Publication No. US20040203103A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/820,536  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: 10/303,128

```

; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
US-10-820-536-12

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60
Db 1 MKKVIIAGNPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
Db 61 TLKHLIQOEYETELIMCSNVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

Qy 121 FHEIFYNQRITSGVVMCAVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVVMCAVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENVVYKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENVVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 9
US-10-845-412-12
; Sequence 12, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-845-412-12

Query Match      100.0%; Score 1563; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60
Db 1 MKKVIIAGNPSLKEIDYSRLPNDVDFRCNQFYEDKYYLGKKCAVFTPGFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120
Db 61 TLKHLIQOEYETELIMCSNVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 120

Qy 121 FHEIFYNQRITSGVVMCAVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVVMCAVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKENVVYKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKENVVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 8
US-10-845-408-12
; Sequence 12, Application US/10845408
; Publication No. US20040203112A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,408
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-845-408-12
```

```
QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPLNSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPLNSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 10
US-10-846-219-12
; Sequence 12, Application US/10846219
; Publication No. US20040219638A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 0196333-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; PRIOR FILING DATE: 2004-05-14
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-846-219-12

Query Match 100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2,1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYLGKKCKAVFYTPGFFFEQY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYLGKKCKAVFYTPGFFFEQY 60

QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120

QY 121 PHEIFYNQRITSGVYMCVAVALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
Db 121 PHEIFYNQRITSGVYMCVAVALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPLNSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPLNSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 11
US-10-821-604-12
; Sequence 12, Application US/10821604
; Publication No. US20040229263A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
```

```
; FILE REFERENCE: 0196333-000111US
; CURRENT APPLICATION NUMBER: US/10/821,604
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-821-604-12

Query Match 100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2,1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYLGKKCKAVFYTPGFFFEQY 60
Db 1 MKKVIIAGNGPSLKEIDYSRLPNDVDFRCNQFYEDKYLGKKCKAVFYTPGFFFEQY 60

QY 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKOLKEFNAYFK 120

QY 121 PHEIFYNQRITSGVYMCVAVALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
Db 121 PHEIFYNQRITSGVYMCVAVALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPLNSNFIIQEKNNYT 240
Db 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLPNSLLANFIELAPLNSNFIIQEKNNYT 240

QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291
Db 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPKKG 291

RESULT 12
US-10-847-983-12
; Sequence 12, Application US/10847983
; Publication No. US20040229272A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 0196333-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; PRIOR FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
```



```
;
;
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-847-983-12

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 13
US-10-821-573-12
; Sequence 12, Application US/10821573
; Publication No. US20040229313A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-821-573-12

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 15
US-10-850-125-12
; Sequence 12, Application US/10850125
; Publication No. US20040259203A1
```

```
;
;
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-847-983-12

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 14
US-10-850-807-12
; Sequence 12, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; PRIOR FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-850-807-12

Query Match      100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60
Db 1 MKKVIIAGNGPSLKEIDYSLRPNDPFDVFRNCQFYEDKYILGKCKAVFTPGFFFEQYY 60

Qy 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120
Db 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFFKQKKEFNAYFK 120

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180
Db 121 FHEIFYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240
Db 181 DRSHYIGHSKNTDIAKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLNSFIIOEKKNYT 240

Qy 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291
Db 241 KDILIPSEAYGKFSKNINFKKIKIKENVYVKLIKDLLRLPSDIKHVFKGK 291

RESULT 15
US-10-850-125-12
; Sequence 12, Application US/10850125
; Publication No. US20040259203A1
```

; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-850-125-12

Query Match 100.0%; Score 1563; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.1e-123;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKVIAGNPSLKEIDYSRLNDFDVFRCNQFYEDKTYLGGKCKAVFYTPGFFPEQYY 60
DB 1 MKKVIAGNPSLKEIDYSRLNDFDVFRCNQFYEDKTYLGGKCKAVFYTPGFFPEQYY 60

QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDPAHLGYDFKQKEFNAYFK 120
DB 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDPAHLGYDFKQKEFNAYFK 120

QY 121 FHEIYFNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180
DB 121 FHEIYFNQRTITSGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180

QY 181 DRSHYIGHSKNTDIKALBFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240
DB 181 DRSHYIGHSKNTDIKALBFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240

QY 241 KDILIPSEAYGKFSKNINFKKIKENVYKLIKDILRLPSDIKHYPK GK 291
DB 241 KDILIPSEAYGKFSKNINFKKIKENVYKLIKDILRLPSDIKHYPK GK 291

Search completed: July 19, 2006, 14:20:08
Job time : 96.0051 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 14:13:57 ; Search time 13.2159 Seconds  
(without alignments)  
1269.788 Million cell updates/sec

Title: US-10-734-719-12

Perfect score: 1563

Sequence: 1 MKKVIIAGNGSLKEIDYSR.....KLIKDLLRLPSGIKHYFGK 291

Scoring table: Blosup62

Gapop 10.0 , Gapext 0.5

Searched: 208217 seqs, 57668156 residues

Total number of hits satisfying chosen parameters: 208217

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.New.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96.5	6.2	496	6	US-10-471-571A-3348 Sequence 3348, Ap
2	95	6.1	701	6	US-10-471-571A-1902 Sequence 1902, Ap
3	92.5	5.9	468	6	US-10-471-571A-5412 Sequence 5412, Ap
4	92.5	5.9	955	6	US-10-471-571A-708 Sequence 708, App
5	92	5.9	396	6	US-10-449-902-50660 Sequence 50660, A
6	92	5.9	512	7	US-11-056-355B-91714 Sequence 91714, A
7	92	5.9	512	7	US-11-056-355B-95470 Sequence 95470, A
8	92	5.9	530	7	US-11-056-355B-91713 Sequence 91713, A
9	92	5.9	530	7	US-11-056-355B-95469 Sequence 95469, A
10	92	5.9	659	7	US-11-056-355B-91712 Sequence 91712, A
11	92	5.9	659	7	US-11-056-355B-95468 Sequence 95468, A
12	91.5	5.9	651	6	US-10-471-571A-252 Sequence 252, App
13	90	5.8	425	6	US-10-953-349-10851 Sequence 10851, A
14	90	5.8	429	6	US-10-953-349-10850 Sequence 10850, A
15	86	5.5	615	7	US-11-056-355B-46743 Sequence 46743, A
16	86	5.5	615	7	US-11-056-355B-48764 Sequence 48764, A
17	86	5.5	676	7	US-11-056-355B-46742 Sequence 46742, A
18	86	5.5	676	7	US-11-056-355B-48763 Sequence 48763, A
19	86	5.5	709	7	US-11-056-355B-46741 Sequence 46741, A
20	86	5.5	709	7	US-11-056-355B-48762 Sequence 48762, A
21	85.5	5.5	437	7	US-11-293-697-4502 Sequence 4502, Ap
22	85.5	5.5	1143	6	US-10-630-629-1 Sequence 1, Appl
23	85	5.4	528	6	US-10-449-902-44124 Sequence 44124, A
24	85	5.4	1337	7	US-11-257-500-29 Sequence 29, Appl
25	84.5	5.4	863	6	US-10-527-411-36 Sequence 36, Appl

26	84.5	5.4	1127	6	US-10-527-411-54 Sequence 54, Appl
27	84.5	5.4	1129	6	US-10-527-411-52 Sequence 52, Appl
28	84	5.4	348	6	US-10-953-349-10852 Sequence 10852, A
29	84	5.4	971	6	US-10-528-563-7 Sequence 7, Appl
30	83.5	5.3	1516	7	US-11-257-500-23 Sequence 23, Appl
31	82.5	5.3	188	6	US-10-471-571A-722 Sequence 722, App
32	82.5	5.3	965	7	US-11-178-560-3 Sequence 3, Appl
33	82	5.2	456	7	US-11-056-355B-79879 Sequence 79879, A
34	82	5.2	478	7	US-11-056-355B-79878 Sequence 79878, A
35	82	5.2	521	7	US-11-056-355B-79877 Sequence 79877, A
36	81.5	5.2	476	6	US-10-449-902-32670 Sequence 32670, A
37	81.5	5.2	965	7	US-11-253-453-3 Sequence 3, Appl
38	81	5.2	877	6	US-10-527-411-157 Sequence 157, App
39	81	5.2	887	6	US-10-527-411-161 Sequence 161, App
40	80.5	5.2	289	6	US-10-471-571A-4924 Sequence 4924, Ap
41	80.5	5.2	863	6	US-10-527-411-38 Sequence 38, Appl
42	80.5	5.2	1127	6	US-10-527-411-50 Sequence 50, Appl
43	80.5	5.2	1129	6	US-10-527-411-48 Sequence 48, Appl
44	80.5	5.2	1217	6	US-10-471-571A-4942 Sequence 4942, Ap
45	80	5.1	488	7	US-11-056-355B-48594 Sequence 48594, A

#### ALIGNMENTS

##### RESULT 1

US-10-471-571A-3348  
; Sequence 3348, Application US/10471571A  
; Publication No. US20060115490A1  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P02692WO  
; CURRENT APPLICATION NUMBER: US/10/471,571A  
; CURRENT FILING DATE: 2003-09-12  
; PRIOR APPLICATION NUMBER: GB-0107661.1  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 5642  
; SOFTWARE: SeqWin99, version 1.03  
; SEQ ID NO 3348  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(496)  
; OTHER INFORMATION: hypothetical protein  
US-10-471-571A-3348

Query Match	6.2%	Score 96.5;	DB 6;	Length 496;
Best Local Similarity	20.3%	Pred. No. 1.3;		
Matches	66;	Conservative 51;	Mismatches 113;	Indels 95; Gaps 17;
QY	12	SLKEIDYRLPNDPFRCNQF----	YFEDKYY-----	LGKKKAVFY-TPGF- 54
Db	90	TLKHVENS---HDFRIYDQERFLMVAHFQDPKYRLDYVNHFDSPQRKVRDFYVGRFL 146		
QY	55	-----FFEQYYTLKHLIQEYETELIMCSNVQAHLENENFVKTF-----	YDFPDDAHLG 105	
Db	147	SCSRILVDKQQTLCFFYFNPEGDTKLEKYSYKDGKPEVQKIIVYANKQYFFNNETELG 206		
QY	106	YDFPKQLKEFNAYKFFHEIYENQRIITSGVYMC-----	AVAIALGYKEIVLSGID--- 154	
Db	207	AFFIKQL-----YHGDLFFSDR---NVYTAIFNLTPESIPIVAVLHSTHAKNIDALD 257		
QY	155	-----FYQNGSSY---AFDTKQENLLKLAPDKNDRSHYI-----	GHSKNNTDI 194	
Db	258	SSPFKNVYKAMENLSRYRAIIVSTEQQKL-----	DVEKRINHTIPWNIPVGYSEITD 312	
QY	195	KALEFLEKTYK-IKLYCLCPNLSLLANFIELAPNLNSNFTIOEKNNYTKDILIPSS- 249		
Db	313	PVQTLDRSVKLSIVARYSPKQLHQQLIELKRLVS-----	YVPKIELHMY 358	

```
Qy 250 AYGFSSKNIN--FKKIKIKENVYK 272
Db 359 GFGSESKLNELIOKYGLENHVYLR 383

RESULT 2
US-10-471-571A-1902
; Sequence 5412, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 1902
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(701)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-1902

Query Match 6.1%; Score 95; DB 6; Length 701;
Best Local Similarity 19.9%; Pred. No. 2.7;
Matches 72; Conservative 45; Mismatches 91; Indels 154; Gaps 19;

Qy 15 EIDYSRLP-----NDPDPV-----RCNQFYFEDKYLGKKCAVFPYTPGPF 57
Db 324 DINEHLKPVPVFDDIHDIISOTNNLNMRCFEKLFERNIGLAIAIKST-----Q 375

Qy 58 QYVTLKHLI-----QNOEYETE-----LIMCSNYNOA-----HLENEN----- 90
Db 376 QFETMKQLLITFLQGNQDYKTSKLVKFMVLVFCNSMTABEHLCHLKIKKNKEIKYSV 435

Qy 91 FVKTFYFDPDAHLYGYPFFKOLKEFNAYF-----KPH--EIVFNQRI 131
Db 436 TVDGFLETYSTVEQVDMVORLK-FHYFYFDIENSKTATHLITKNQHYQDTDFEQ--- 491

Qy 132 SGVVMCAVAIALGYKEIYL-SGIDFYQ-----NGSSYAFDTKQENLLK----- 173
Db 492 -----YKKFILDSGISSTQFVNNLSVSGFKYTNDSGNPIQLSDIVYHLIAL 538

Qy 174 -----LAPDFKNDRSHYIG-HSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFELAPN 226
Db 539 LRYGGISYQLLDDHSNYSILNKGSGPLPLMHLKMF-----PFVNEDIE 585

Qy 227 LNSNFIQEKNNYTKOILIPSSSEAYGKFSKNINFKIKIKENVYKIKDLRLPSDIKH 286
Db 586 ITNNVLSRKDN-----NYHF-----LLFNKIND--RYMSDVQK 617

Qy 287 YF 288
Db 618 DF 619

RESULT 3
US-10-471-571A-5412
; Sequence 5412, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
```

```
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 5412
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-5412

Query Match 5.9%; Score 92.5; DB 6; Length 468;
Best Local Similarity 22.5%; Pred. No. 2.7;
Matches 62; Conservative 46; Mismatches 78; Indels 89; Gaps 18;

Qy 56 FEQYVTLKHLIQOEYETELIMCSNYNOAHLENENFVKTFYDYFPPDAH--LGYDFFKQLK 113
Db 234 FEQY--KKNVLDKY-----YSEVYKSDNLIKVFSE--DVMNIMGY-----K 273

Qy 114 EFNAYF-----KPEHIYFNQRI-TSGVVMCAVAIALGYKEIYLSGIDFYONGSSYAF 164
Db 274 KYNEYFTNRDDTTLTITNLVYIQSGFSGTEGNNRVIDRIGKHEKL-----QEIMTYQP 325

Qy 165 DTQOENLL--KLAPDFKNDRSHYI--GHSKNTDIKALEFLEKTYKIKLYC-LCPNSILLAN 219
Db 326 ENRKTVLIFYKIQPD-----NIYITNQYKQEAVSRLKRILEQSIIPNKEYIDISEISLITF 380

Qy 220 FIELAPNLSNFIQEKNNYTKOILIPSSSEAYGKFSKN-----INFKKIKIKENVYKYL 273
Db 381 F---KPEYNKQHVNEEFS-----LSSSKIKNLAKENDGIILYNLFK-----DVNYQL 425

Qy 274 I-----KOLLRLPS-DIKHYFK 289
Db 426 ISEGNTFFSAKPYLHCVFNKDIANMDKVDIENFFK 460

RESULT 4
US-10-471-571A-708
; Sequence 708, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 708
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-471-571A-708

Query Match 5.9%; Score 92.5; DB 6; Length 955;
Best Local Similarity 20.7%; Pred. No. 6.8;
Matches 56; Conservative 44; Mismatches 72; Indels 99; Gaps 14;

Qy 62 LKHLIQOEYETELIMCSNYNOAHLENENFVKTFYDYFPPDAHLYGYPFFKOLKEFNAYFKF 121
Db 36 LKYLIRMSRTPPTPGLLSGINLGHFVNEP-----TRLKVGNSIQKY 76

Qy 122 HEIYFNQRIITSGVVMCAVAIALGYKEI-YLSGID-FYONGSSYAFDTKQENLLKLAPDFK 179
Db 77 VKV-----DGEWL-----YKLVSYIESIDEYYQN-----LKVITWNSK 108

Qy 180 -----NDRSH-----YIGHSKNTDI-----KALEFLEKTYKIKLYCLCPNSLLANFIEL 223
Db 109 AHIINDRIYLNQSAIYLNKKDTSFSIKNSSELLVFIKT-----VTNNNI--TFSNL 159

Qy 224 APNLSNFIQOE-----KNYTKOILIPSSSEAYGKFSKNINF--KKIKIKENVYKYL 273
Db 160 AEKINQPEINDITKVKVYIHNLSKSEIISTIRPPLSYSDNLSNLTILNKLHNDDFVKK 219
```

QY 274 IKDLRL-----LPSDIKHVFK 289  
Db 220 IREIQKILAYEKTEIGFGEELYKDIHHMK 250

## RESULT 5

US-10-449-902-50660  
; Sequence 50660, Application US/10449902  
; Publication No. US20060123505A1  
; GENERAL INFORMATION:  
; APPLICANT: National Institute of Agricultural Sciences.  
; APPLICANT: Bio-Oriented Technology Research Advancement Institution.  
; APPLICANT: The Institute of Physical and Chemical Research.  
; APPLICANT: Foundation for Advancement of International Science.  
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF  
; FILE REFERENCE: MOA-A0205Y1-US  
; CURRENT APPLICATION NUMBER: US/10/449,902  
; CURRENT FILING DATE: 2003-05-29  
; PRIOR APPLICATION NUMBER: JP 2002-203269  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: JP 2002-383870  
; PRIOR FILING DATE: 2002-12-11  
; NUMBER OF SEQ ID NOS: 56791  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 50660  
; LENGTH: 396  
; TYPE: prt  
; ORGANISM: Oryza sativa  
US-10-449-902-50660

Query Match 5.9%; Score 92; DB 6; Length 396;  
Best Local Similarity 19.3%; Pred. No. 2.4;  
Matches 60; Conservative 58; Mismatches 97; Indels 96; Gaps 19;

QY 9 NGPSLKEIDYRLNDFVFCNQFYEDKYLKCKKAVYTPGFFPEQY--TLKHILI 66  
Db 119 NGRKLDLDAKL-----KEKY-----PESILYKGYF--QYEDSLEWYF 157  
QY 67 QNOEY-----TELIMCSYNOAHLENENFVKTFYDFPDALGY--DFFKOLKEFNAY 118  
Db 158 DPERFQPAALDNYQELVLC--NGLYMDWDQYHSNYTY--ESDLAYVKFCELAHKTW 213  
QY 119 FKFHEIYF--NQRITSG-----VYMCVAIAIALGYKEI-----YLSGIDF-- 155  
Db 214 FQDYLVLIAVEDKITMGQWDKNTVYLQAMKIALIRVSLMVMTAFQYIWSMRD 273  
QY 156 --YONGSSYADTYQENLLKLPDFKNDRSHVIGHS-----KNTDIK-----ALEPTE 201  
Db 274 CNYKDFDGVYFEV--WKRVAKQKMEFTDALSLEHREDMFPLRNVDIKDLRTRGRFRGMK 332  
QY 202 KTYKIKLYCLCPNSILANFIELAPNLNSNFIIOEKNVTKDI---LIPSSAYGKFSKNI 258  
Db 333 ENYDLVYACI-----DETVE-----KEARLIKDAIEMIPKQTYLDIARN-- 375  
QY 259 NFKKIKIKENV 269  
Db 376 ---KIKIAEDI 383

## RESULT 6

US-11-056-355B-91714  
; Sequence 91714, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190

; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 91714  
; LENGTH: 512  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(512)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728960  
US-11-056-355B-91714

Query Match 5.9%; Score 92; DB 7; Length 512;  
Best Local Similarity 21.4%; Pred. No. 3.3;  
Matches 75; Conservative 50; Mismatches 126; Indels 100; Gaps 18;  
QY 5 IIAAGN-GPSLKEIDYSRL-----PND-----FDVFRCNQFYEDKYLKCKKAVFYTP 52  
Db 114 LIAGRVDEALKQLEMMVRKLNPNNEATIRTFVHGIFRCL-----PPCKAFEVLV 162  
QY 53 GFFPEQYTLKHLIQNOYETELIMCSYNOAHLENENFVKTFD--YFPDAHLGYDFFK 110  
Db 163 GFMEKD-----SNLQRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERYIFDS----- 208  
QY 111 QLKEFNA---YFKFHEIYFNORITSGV-----YMCVAIAIAL-----GYKEI 148  
Db 209 --STFNAAMSCLLKGHLVETCTRIFDGFVSRGVKPGFNGYLVLVQALLNAQRFSEGDRL 266  
QY 149 YLSGIDFYQNGSSYADTYQENLLK-----LAPDFKNDRSHYIGHSK 190  
Db 267 KQMGVDGLLS--SVYSYNAVIDCLCKARRIENAMFTEMQDRGISPNLVTFTFLSGYSV 325  
QY 191 NTDIKALE--FLEK-----TYKIKLYCLCPNSILANFIELAPNLNSNFIIOEKN 238  
Db 326 RGDVKVKGHVLEKLLVHGFKPDVITFSLIINCLCRAKEIKDAPDCFKEMLEWGI--EPNE 383  
QY 239 YTKDILIPSSAYGKFSKNNIFKKIKIKENVYKLIKDLRLPSDIKHVFK 289  
Db 384 ITYNILIRSCCTGTDTRSVKL--FAKMKEN---GLSPDLVAYNATIOQSFCK 430

## RESULT 7

US-11-056-355B-95470  
; Sequence 95470, Application US/11056355B  
; Publication No. US20060150283A1  
; GENERAL INFORMATION:  
; APPLICANT: Brover, Vyacheslav  
; APPLICANT: Alexandrov, Nikolai  
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: 2750-1590PUS2  
; CURRENT APPLICATION NUMBER: US/11/056,355B  
; CURRENT FILING DATE: 2005-02-14  
; PRIOR APPLICATION NUMBER: 60/544,190  
; PRIOR FILING DATE: 2004-02-13  
; NUMBER OF SEQ ID NOS: 119966  
; SEQ ID NO 95470  
; LENGTH: 512  
; TYPE: prt  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: (1)..(512)  
; OTHER INFORMATION: Ceres Seq. ID no. 12728960  
US-11-056-355B-95470

Query Match 5.9%; Score 92; DB 7; Length 512;  
Best Local Similarity 21.4%; Pred. No. 3.3;  
Matches 75; Conservative 50; Mismatches 126; Indels 100; Gaps 18;  
QY 5 IIAAGN-GPSLKEIDYSRL-----PND-----FDVFRCNQFYEDKYLKCKKAVFYTP 52  
Db 114 LIAGRVDEALKQLEMMVRKLNPNNEATIRTFVHGIFRCL-----PPCKAFEVLV 162



```

; LENGTH: 659
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: (1)..(659)
; OTHER INFORMATION: Ceres Seq. ID no. 12728958
US-11-056-355B-91712

Query Match      5.9%; Score 92; DB 7; Length 659;
Best Local Similarity 21.4%; Pred. No. 4.6;
Matches 75; Conservative 50; Mismatches 126; Indels 100; Gaps 18;

QY 5 IIAAG-GPSLKEIDYSLR-----PND-----FDVFRCNQFYFEDKYVGLKKCAVFTTP 52
Db 261 LIAGRVDEALKQLEMVRKLNPNENATIRTFVHGIFRCL-----PPCKAFEVLV 309
QY 53 GFFPEQYVTLKHLIQNOEYETELIMCSNNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110
Db 310 GFMEKD-----SNLQRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERYIPDS----- 355
QY 111 QLKKEFNA-----YFKPHEIYFNQRTITSGV-----YMCVAIAIAL-----GYKEI 148
Db 356 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVVQALLNAQRFSEGDRL 413
QY 149 YLSGIDFYONGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190
Db 414 KOMGVGGLLS-SVSYNAVIDCLCKARRIENAAAMFLTEMQDRGISPNLVTFTFLSGYSV 472
QY 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNN 238
Db 473 RGVKVKHGVLEKLLVHGFKPDVITFSLIINCLCKRAKEIKDAFDCFEMLEWGI--EPNE 530
QY 239 YTKDILIPSSBAYGKFSKNINFKIKENVYKLIKDLRLPSDIKHYPK 289
Db 531 ITYNILIRSCCTGDTDRSVKL-FAKMKEN---GLSPDLVAYNATIOQSCK 577

RESULT 11
US-11-056-355B-95468
; Sequence 95468, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 95468
; LENGTH: 659
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: (1)..(659)
; OTHER INFORMATION: Ceres Seq. ID no. 12728958
US-11-056-355B-95468

Query Match      5.9%; Score 92; DB 7; Length 659;
Best Local Similarity 21.4%; Pred. No. 4.6;
Matches 75; Conservative 50; Mismatches 126; Indels 100; Gaps 18;

QY 5 IIAAG-GPSLKEIDYSLR-----PND-----FDVFRCNQFYFEDKYVGLKKCAVFTTP 52
Db 261 LIAGRVDEALKQLEMVRKLNPNENATIRTFVHGIFRCL-----PPCKAFEVLV 309
QY 53 GFFPEQYVTLKHLIQNOEYETELIMCSNNOAHLENENFVKTFYD--YFPDAHLGYDFFK 110

```

```

Db 310 GFMEKD-----SNLQRVGYDAVLYCLSNNSMAK-ETGQFLRKIGERYIPDS----- 355
QY 111 QLKKEFNA-----YFKPHEIYFNQRTITSGV-----YMCVAIAIAL-----GYKEI 148
Db 356 --STFNAAMSCLLKGHDLVETCRIFDGFVSRGVKPGFNGYLVVQALLNAQRFSEGDRL 413
QY 149 YLSGIDFYONGSSYAFDTKQENLLK-----LAPDFKNDRSHYIGHSK 190
Db 414 KOMGVGGLLS-SVSYNAVIDCLCKARRIENAAAMFLTEMQDRGISPNLVTFTFLSGYSV 472
QY 191 NTDIKALE-FLEK-----TYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNN 238
Db 473 RGVKVKHGVLEKLLVHGFKPDVITFSLIINCLCKRAKEIKDAFDCFEMLEWGI--EPNE 530
QY 239 YTKDILIPSSBAYGKFSKNINFKIKENVYKLIKDLRLPSDIKHYPK 289
Db 531 ITYNILIRSCCTGDTDRSVKL-FAKMKEN---GLSPDLVAYNATIOQSCK 577

RESULT 12
US-10-471-571A-252
; Sequence 252, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 252
; LENGTH: 651
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(651)
; OTHER INFORMATION: hypothetical ABC
US-10-471-571A-252

Query Match      5.9%; Score 91.5; DB 6; Length 651;
Best Local Similarity 22.4%; Pred. No. 5;
Matches 50; Conservative 29; Mismatches 67; Indels 77; Gaps 11;

QY 14 KEIDY-SRLPNDP--DVFR-----CNQFYFEDKYVGLKKCAVFTTPGFFEQY-- 59
Db 112 EETDWLSKHANDYSDTYKTHMSRYESLSNQFLEQYQVESKIKTVLYGLNFSSEDFNK 171
QY 60 -----YTLKHLIQNOEYETELIMCSNNOAHLENENFVKTFDYFPDAHLGYD 107
Db 172 PINDFSGGQKTRLSLAQMLN---EPDLLLDE-----PTNHLDE 209
QY 108 FFKOLKEFNAYFK-----FHEIYFNQRTITSGVMCAVAIALGYKEIYLSGI-DFYQNGS 160
Db 210 TTKWLEDYLRYPKGAIVIIISHDRYFLDKIVTQIY-----DVALGVKRYGVNVEEFTQORD 265
QY 161 SYAFDTKQENLLKLAAPDFKNDRSHYIGHSKNTDIKALE-FLEK 202
Db 266 LYYQKRMQEY-----ESQQAIEIKLETVEK 291

RESULT 13
US-10-953-349-10851
; Sequence 10851, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2

```

; CURRENT APPLICATION NUMBER: US/10/953,349  
 ; CURRENT FILING DATE: 2004-09-30  
 ; NUMBER OF SEQ ID NOS: 40252  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 10851  
 ; LENGTH: 425  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 US-10-953-349-10851

Query Match 5.8%; Score 90; DB 6; Length 425;  
 Best Local Similarity 18.9%; Pred. No. 3.9; Indels 112; Gaps 20;  
 Matches 68; Conservative 69; Mismatches 111;

```

QY 11 PSLKEIDYSLRPNDFVRCNQFYEDK-----YVLGKKCAVFTPGPFF 56
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 76 PTMEETYSVYHSPEAFRLNVAEMKRFKVIYIPDGPNTFYOTPRKVTGYASEGYFF 135
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 57 E-----QYITLK-----HLIQNO--EYETELIMCSYNOAHLENENFVKTFY 96
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 136 QNIRESRFTLDPDEADLFFPISCHKMRGKGTSYENMTVIQYVVDGLIAK-----Y 188
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 97 DYFDDAHLGYDFKQLKEFNAYRFHEIYFNQRTSGV-----YMCVAIALGY-- 145
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 189 PYW-NRTLGAHF-----FVTCADV--GVRAFEGLLIKNTIRVVGSPSYNVGFIP 237
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 146 -KEIYL-----SGIDFYQNGSSYAFDTKQEN---LLKLAPDFKNDRSHYIGHSK- 190
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 238 HKDVALPQVLQPPALPAGNDVNRITLGFAGHNSKIRVILAHVWENDTDLDSNNRI 297
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 191 NTDIKALEFLEKTYKIKLYCLCP-----NS-----LLANFIELAPNLSNF 231
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 298 NRATGHLVYQKRFYRTK-FCICPGGSQVNSARITDSIHYGCIPIVLSYDYLDPNDILNW 356
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 232 -----IIQEKNYT-KDIL--IPSEAYG-----KFSKNINFKKIKIKENVYVKIKDL 277
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 357 RKFAVLREQDVYLNKQLLNKNIHSEFVSLHNNLVKQKHFWNSPPVKFADFAMIMYEL 416
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
  
```

RESULT 14  
 US-10-953-349-10850  
 ; Sequence 10850, Application US/10953349  
 ; Publication No. US20060107345A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ALEXANDROV, Nikolai et al.  
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
 ; FILE REFERENCE: 2750-1579FUS2  
 ; CURRENT APPLICATION NUMBER: US/10/953,349  
 ; CURRENT FILING DATE: 2004-09-30  
 ; NUMBER OF SEQ ID NOS: 40252  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 10850  
 ; LENGTH: 429  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 US-10-953-349-10850

Query Match 5.8%; Score 90; DB 6; Length 429;  
 Best Local Similarity 18.9%; Pred. No. 3.9;  
 Matches 68; Conservative 69; Mismatches 111; Indels 112; Gaps 20;

```

QY 11 PSLKEIDYSLRPNDFVRCNQFYEDK-----YVLGKKCAVFTPGPFF 56
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 80 PTMEETYSVYHSPEAFRLNVAEMKRFKVIYIPDGPNTFYOTPRKVTGYASEGYFF 139
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 57 E-----QYITLK-----HLIQNO--EYETELIMCSYNOAHLENENFVKTFY 96
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 140 QNIRESRFTLDPDEADLFFPISCHKMRGKGTSYENMTVIQYVVDGLIAK-----Y 192
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 97 DYFDDAHLGYDFKQLKEFNAYRFHEIYFNQRTSGV-----YMCVAIALGY-- 145
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 193 PYW-NRTLGAHF-----FVTCADV--GVRAFEGLLIKNTIRVVGSPSYNVGFIP 241
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
  
```

```

QY 146 -KEIYL-----SGIDFYQNGSSYAFDTKQEN---LLKLAPDFKNDRSHYIGHSK- 190
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 242 HKDVALPQVLQPPALPAGNDVNRITLGFAGHNSKIRVILAHVWENDTDLDSNNRI 301
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 191 NTDIKALEFLEKTYKIKLYCLCP-----NS-----LLANFIELAPNLSNF 231
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 302 NRATGHLVYQKRFYRTK-FCICPGGSQVNSARITDSIHYGCIPIVLSYDYLDPNDILNW 360
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 232 -----IIQEKNYT-KDIL--IPSEAYG-----KFSKNINFKKIKIKENVYVKIKDL 277
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 361 RKFAVLREQDVYLNKQLLNKNIHSEFVSLHNNLVKQKHFWNSPPVKFADFAMIMYEL 420
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
  
```

RESULT 15  
 US-11-056-355B-46743  
 ; Sequence 46743, Application US/11056355B  
 ; Publication No. US20060150283A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brover, Vyacheslav  
 ; APPLICANT: Alexandrov, Nikolai  
 ; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding  
 ; FILE REFERENCE: 2750-1590FUS2  
 ; CURRENT APPLICATION NUMBER: US/11/056,355B  
 ; CURRENT FILING DATE: 2005-02-14  
 ; PRIOR APPLICATION NUMBER: 60/544,190  
 ; PRIOR FILING DATE: 2004-02-13  
 ; NUMBER OF SEQ ID NOS: 119966  
 ; SEQ ID NO 46743  
 ; LENGTH: 615  
 ; TYPE: PRT  
 ; ORGANISM: Arabidopsis thaliana  
 ; NAME/KEY: peptide  
 ; LOCATION: (1)..(615)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 12600827  
 US-11-056-355B-46743

Query Match 5.5%; Score 86; DB 7; Length 615;  
 Best Local Similarity 23.6%; Pred. No. 14;  
 Matches 52; Conservative 35; Mismatches 85; Indels 48; Gaps 12;

```

QY 12 SLKEIDYSLRPNDFVRCNQFYEDKYY-----LQKKCAVFTPGPFF-----EQ 58
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 193 SLQRYVIGISYDQQLKKYLOFLEAKYDRLGQKQELFFSHQLSPGSYFFPLGLTRV 252
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 59 YITLKHILIQOEYE---TELIMCSYNOAHLENENFVKTFYDYPDAHLGYDFFKQ---L 112
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 253 YNRLMDFIKQYWHRGYTEVITPNMNMELWQTSCHA----DNYKDNMTFENIEKQEFGL 308
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 113 KEFNAYKFHEIYFNQRTSGVYMCVAIALGYKEIYLSGIDF---YONGSSYAFDTKQE 169
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 309 KPMNC--PGHCLIFQHRVRS-----YELPMRLADFGVLHNEASGALSG--- 351
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 170 NLLKLAPDFKNDRSHYIGHSKNT--DIK-ALEFLEKTYKI 206
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 352 --LTRVRFRQDDAHIFCTTEQVKGVEGVLEFIDYVYKV 389
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
  
```

Search completed: July 19, 2006, 14:21:08  
 Job time : 14.2159 secs



GenCore version 5.1.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:55:46 ; Search time 5050.55 Seconds  
(without alignments)  
11053.473 Million cell updates/sec

Title: US-10-734-719-13

Perfect score: 873

Sequence: 1 atgaaaaaagtattattgc.....agcattatttcaaggaaaa 873

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_env.\*  
2: gb\_pat.\*  
3: gb\_ph.\*  
4: gb\_pl.\*  
5: gb\_pr.\*  
6: gb\_ro.\*  
7: gb\_scs.\*  
8: gb\_sy.\*  
9: gb\_un.\*  
10: gb\_vi.\*  
11: gb\_ov.\*  
12: gb\_htg.\*  
13: gb\_in.\*  
14: gb\_cm.\*  
15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	873	100.0	873	2	AX934436 Sequence
2	873	100.0	873	2	AR481787 Sequence
3	873	100.0	873	2	AR527386 Sequence
4	873	100.0	873	2	AR609665 Sequence
5	873	100.0	873	2	AR689943 Sequence
6	873	100.0	873	2	AR691839 Sequence
7	873	100.0	12390	15	AF401528 Campyloba
8	871.4	99.8	4749	15	AF305571 Campyloba
9	869.8	99.6	876	2	AX934431 Sequence
10	869.8	99.6	876	2	BD249794 Campyloba
11	869.8	99.6	876	2	AR271703 Sequence
12	869.8	99.6	876	2	AR481785 Sequence
13	869.8	99.6	876	2	AR527384 Sequence
14	869.8	99.6	876	2	AR609663 Sequence
15	869.8	99.6	876	2	AR689941 Sequence
16	869.8	99.6	876	2	AR691837 Sequence
17	869.8	99.6	6047	15	AY661458 Campyloba
18	869.8	99.6	24425	15	AY422197 Campyloba

19	869.8	99.6	24437	15	AF167344	AF167344 Campyloba
20	866.6	99.3	873	2	AX934434	AX934434 Sequence
21	866.6	99.3	873	2	AR481786	AR481786 Sequence
22	866.6	99.3	873	2	AR527385	AR527385 Sequence
23	866.6	99.3	873	2	AR609664	AR609664 Sequence
24	866.6	99.3	873	2	AR689942	AR689942 Sequence
25	866.6	99.3	873	2	AR691838	AR691838 Sequence
26	866.6	99.3	11474	15	AF215659	AF215659 Campyloba
27	852.4	97.6	1303	15	AY681242	AY681242 Campyloba
28	841	96.3	876	2	AX934427	AX934427 Sequence
29	841	96.3	876	2	BD249792	BD249792 Campyloba
30	841	96.3	876	2	AR271701	AR271701 Sequence
31	841	96.3	876	2	AR481783	AR481783 Sequence
32	841	96.3	876	2	AR527382	AR527382 Sequence
33	841	96.3	876	2	AR609661	AR609661 Sequence
34	841	96.3	876	2	AR689939	AR689939 Sequence
35	841	96.3	876	2	AR691835	AR691835 Sequence
36	839.4	96.2	11442	15	AF400048	AF400048 Campyloba
37	834.6	95.6	876	15	AF216647	AF216647 Campyloba
38	834.6	95.6	6047	15	AY644679	AY644679 Campyloba
39	834.6	95.6	11474	2	AX934424	AX934424 Sequence
40	834.6	95.6	11474	2	BD249790	BD249790 Campyloba
41	834.6	95.6	11474	2	AR271699	AR271699 Sequence
42	834.6	95.6	11474	2	AR481781	AR481781 Sequence
43	834.6	95.6	11474	2	AR527380	AR527380 Sequence
44	834.6	95.6	11474	2	AR609659	AR609659 Sequence
45	834.6	95.6	11474	2	AR689937	AR689937 Sequence

ALIGNMENTS

RESULT 1  
AX934436  
LOCUS AX934436 873 bp DNA linear PAT 05-JAN-2004  
DEFINITION Sequence 13 from Patent WO02074942.  
ACCESSION AX934436  
VERSION AX934436.1 GI:40641682  
KEYWORDS Campylobacter jejuni  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.  
REFERENCE 1  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics  
JOURNAL Patent: WO 02074942-A 13 26-SEP-2002;  
National Research Council of Canada (CA)  
FEATURES  
Location/Qualifiers  
source  
1..873  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1..>873  
/note="unnamed protein product; Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C. jejuni O:36"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04171.1"  
/db\_xref="GI:40641683"  
/translation="MKKVIILAGPSLSKEIDYSLRPNDFVFCNQFYEDKDYLGKK CKTYPTNFPEQYITLUHLIQNOEYETELIMCSNNOAHLENENFVKTFYDFPDA HLGVDFFKQLKEFNAPDFEYFNQITSGVYMCVAVALGEYKLYSLGIDFYQNG SYADFTKQENLLKAPDFKNDKSHYIGHSKNTDIIKALEFKYKLYKLCIPNSLIA NFIELAPLNSNFIIOEKNNYTKDILIPSEAYGPKSKNFKKIKIKENVYKLIKD LLRLPSDIKHYFKGK"

Query Match 100.0%; Score 873; DB 2; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.3e-214;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAACTAT 120  
QY 121 CTGTAAGAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
DB 121 CTGTAAGAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 240  
DB 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 240  
QY 241 TACCAACAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTAAACCTTTTACGATTTTCC 300  
DB 241 TACCAACAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTAAACCTTTTACGATTTTCC 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 360  
DB 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 360  
QY 361 TTTCAGCAATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 420  
DB 361 TTTCAGCAATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 540  
DB 481 TCTTATGCTTTTGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 540  
QY 541 GATGCTCACTATATCGGACATAGTAAAGAAATACAGATATAAGAGCTTTAGAAATTTCTA 600  
DB 541 GATGCTCACTATATCGGACATAGTAAAGAAATACAGATATAAGAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTTAAAGAAATGTTTATACAGTTGATAAGAGCTTTTAAAGATTA 840  
DB 601 GAAAAAATTTTAAAGAAATGTTTATACAGTTGATAAGAGCTTTTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

RESULT 2  
AR481787  
LOCUS AR481787 873 bp DNA linear PAT 14-MAY-2004  
DEFINITION Sequence 13 from patent US 6699705.  
ACCESSION AR481787  
VERSION AR481787.1 GI:47243422  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 873)

AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
ORIGIN

Gilbert, M. and Wakarchuk, W.W.  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
Patent: US 6699705-A 13 02-MAR-2004;  
National Research Council of Canada; Ottawa;  
CAX;  
Location/Qualifiers  
1. .873  
/organism="unknown"  
/mol\_type="genomic DNA"

Query Match 100.0%; Score 873; DB 2; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.3e-214;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAACTAT 120  
DB 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAACTAT 120  
QY 121 CTGTAAGAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
DB 121 CTGTAAGAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 240  
DB 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 240  
QY 241 TACCAACAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTAAACCTTTTACGATTTTCC 300  
DB 241 TACCAACAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTAAACCTTTTACGATTTTCC 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 360  
DB 301 GATGCTCAATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 360  
QY 361 TTTCAGCAATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 420  
DB 361 TTTCAGCAATTTTAAATCCAAATCAAGATATGAGACCGAATTAATGCTTTTAA 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 540  
DB 481 TCTTATGCTTTTGGATATGATTTTAAACCACTTAAAGAAATTAATGCTTTTAA 540  
QY 541 GATGCTCACTATATCGGACATAGTAAAGAAATACAGATATAAGAGCTTTAGAAATTTCTA 600  
DB 541 GATGCTCACTATATCGGACATAGTAAAGAAATACAGATATAAGAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTTAAAGAAATGTTTATACAGTTGATAAGAGCTTTTAAAGATTA 840  
DB 601 GAAAAAATTTTAAAGAAATGTTTATACAGTTGATAAGAGCTTTTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

```
Db      841 CCTAGTGATATAAGCAATTAATTCAGAGGAAA 873
|||||
AR527386
LOCUS      AR527386                873 bp    DNA        linear    PAT 08-OCT-2004
DEFINITION Sequence 13 from patent US 6723545.
ACCESSION AR527386
VERSION   AR527386.1 GI:53914407
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE  1 (bases 1 to 873)
AUTHORS   Gilbert, M. and Wakarchuk, W.W.
TITLE     Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL   Patent: US 6723545-A 13 20-APR-2004;
          National Research Council of Canada; Ottawa;
          CAX;

FEATURES   Location/Qualifiers
            source
              1..873
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 873; DB 2; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.3e-214;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAATTTAAAGAAATTTGATTATTCAGG 60
Db      1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAATTTAAAGAAATTTGATTATTCAGG 60

Qy      61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db      61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy      121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTTTGGACAAATACTAC 180
Db      121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTTTGGACAAATACTAC 180

Qy      181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTTAAT 240
Db      181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTTAAT 240

Qy      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCT 300
Db      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCT 300

Qy      301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db      301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy      361 TTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTCAGTAGGCC 420
Db      361 TTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTCAGTAGGCC 420

Qy      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy      481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db      481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy      541 GATGCTCACA CTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db      541 GATGCTCACA CTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy      601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAATAGTCTTTTACGAAATTT 660
Db      601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAATAGTCTTTTACGAAATTT 660

Qy      661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACTACT 720
Db      661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACTACT 720

Qy      721 AAAGATATATCTCATACCTCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db      721 AAAGATATATCTCATACCTCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780

Qy      781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Db      781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840

Qy      841 CCTAGTGATATAAGCAATTAATTCAGAGGAAA 873
|||||
AR609665
LOCUS      AR609665                873 bp    DNA        linear    PAT 15-DEC-2004
DEFINITION Sequence 13 from patent US 6825019.
ACCESSION AR609665
VERSION   AR609665.1 GI:56664965
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE  1 (bases 1 to 873)
AUTHORS   Gilbert, M. and Wakarchuk, W.W.
TITLE     Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL   Patent: US 6825019-A 13 30-NOV-2004;
          National Research Council of Canada; Ottawa;
          CAX;

FEATURES   Location/Qualifiers
            source
              1..873
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 873; DB 2; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.3e-214;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db      1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60

Qy      61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db      61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Qy      121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTTTGGACAAATACTAC 180
Db      121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTTTGGACAAATACTAC 180

Qy      181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTTAAT 240
Db      181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAATTAATATGTTCTTAAT 240

Qy      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCT 300
Db      241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCT 300

Qy      301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db      301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy      361 TTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTCAGTAGGCC 420
Db      361 TTTCAGAAATTTATTTCAATCAAAGAAATACCTCAGGGGTCTATATGTCAGTAGGCC 420

Qy      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy      481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db      481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy      541 GATGCTCACA CTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db      541 GATGCTCACA CTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy      601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAATAGTCTTTTACGAAATTT 660
Db      601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAATAGTCTTTTACGAAATTT 660

Qy      661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACTACT 720
Db      661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACTACT 720

Qy      721 AAAGATATATCTCATACCTCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
Db      721 AAAGATATATCTCATACCTCTCTAGTAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780

Qy      781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Db      781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840

Qy      841 CCTAGTGATATAAGCAATTAATTCAGAGGAAA 873
|||||
AR609665
LOCUS      AR609665                873 bp    DNA        linear    PAT 15-DEC-2004
DEFINITION Sequence 13 from patent US 6825019.
ACCESSION AR609665
VERSION   AR609665.1 GI:56664965
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE  1 (bases 1 to 873)
AUTHORS   Gilbert, M. and Wakarchuk, W.W.
TITLE     Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL   Patent: US 6825019-A 13 30-NOV-2004;
          National Research Council of Canada; Ottawa;
          CAX;

FEATURES   Location/Qualifiers
            source
              1..873
              /organism="unknown"
              /mol_type="genomic DNA"

ORIGIN
```



```
QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTGATTTCACGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTGATTTCACGG 60
QY 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACACTAC 180
Db 121 CTTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACACTAC 180
QY 181 ACTTTTAAACATTTAATCAAAATCAAGATATGAGCGGACTAATATGTTCTTAAT 240
Db 181 ACTTTTAAACATTTAATCAAAATCAAGATATGAGCGGACTAATATGTTCTTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300
QY 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360
Db 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360
QY 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
QY 541 GATGCTCACAATATTCGGACATAGTAAATAACAGATATAAAAGCTTTTGAATTTCTA 600
Db 541 GATGCTCACAATATTCGGACATAGTAAATAACAGATATAAAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTCAAAATAAAACTATATGCTTATGCTTCTTAATAGTCTTTTAGCAATTTT 660
Db 601 GAAAAAATTTCAAAATAAAACTATATGCTTATGCTTCTTAATAGTCTTTTAGCAATTTT 660
QY 661 ATAGAATAGCGCCAAATTTAAATTTCAATTTTATCACAAGAAAAATAAATACTACACT 720
Db 661 ATAGAATAGCGCCAAATTTAAATTTCAATTTTATCACAAGAAAAATAAATACTACACT 720
QY 721 AAAGATATCTATACCTTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTAATTT 780
Db 721 AAAGATATCTATACCTTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTAATTT 780
QY 781 AAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
```

## RESULT 7

AF401528

LOCUS

DEFINITION

AF401528 12390 bp DNA linear BCT 03-JAN-2002  
Campylobacter jejuni O:36 heptosyltransferase I (waaC) gene,  
partial cds; lipid A biosynthesis acyltransferase (htrB), putative  
two-domain glycosyltransferase, putative glycosyltransferase,  
truncated beta-1,4-N-acetylglucosaminyltransferase (cgtA-I),  
beta-1,3-galactosyltransferase (cgtB), alpha-2,3-sialyltransferase  
(cst-II), sialic acid synthase (neuB), putative  
N-acetylglucosamine-6-phosphate 2-epimerase (neuC),  
beta-1,4-N-acetylglucosaminyltransferase (cgtA-II), CMP-Neu5Ac  
synthetase (neuA), putative acetyltransferase, and putative

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE

JOURNAL  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

FEATURES  
source

gene

CDS

gene

CDS

CDS

glycosyltransferase (waaV) genes, complete cds; and  
heptosyltransferase II (waaF) gene, partial cds.

AF401528  
GI:15718482

Campylobacter jejuni  
Campylobacter jejuni

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

1 (bases 1 to 12390)

Gilbert M., Karwaski M.F., Bernatchez S., Young N.M., Taboada E.,  
Michniewicz J., Cunningham A.M. and Wakarchuk W.W.

The Genetic Bases for the Variation in the Lipid-Oligosaccharide of  
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF  
SIALYLATED GLANGIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE  
J. Biol. Chem. 277 (1), 327-337 (2002)

11689567

2 (bases 1 to 12390)

Gilbert M., Michniewicz J., Karwaski M.-F., Cunningham A. and  
Wakarchuk W.W.

Direct Submission

Submitted (22-JUL-2001) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada

Location/Qualifiers

1..12390

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strain="ATCC 43456"

/serotype="O:36"

/db\_xref="ATCC:43456"

/db\_xref="taxon:197"

<1..348

/gene="waaC"

<1..348

/gene="waaC"

/codon\_start=1

/transl\_table=11

/product="heptosyltransferase I"

/protein\_id="AAL05984.1"

/db\_xref="GI:15718483"

/translation="FAKEVNLISGIDETHIELAPKNLEELMAFTKMDLIIGNDSP  
THLAFALNKASITIFGATPSYRNAFTINKIIDTGKIQNAKHIDKSDFCITRIEE  
DIFLAKGLLNEK"

338..1225

/gene="htrB"

338..1225

/gene="htrB"

/codon\_start=1

/transl\_table=11

/product="lipid A biosynthesis acyltransferase"

/protein\_id="AAL05985.1"

/db\_xref="GI:15718484"

/translation="MKNSDRIYLSYILKFFVTMPDCILHFLALIVARIAPHLNKK  
HRKINTNLQICFPQVTKERDKSLKIVENFAQFGIDCQNTQNTKEKILNKVFN  
ENFLDALAKRPITFTAHYGNWEILSLAYAKYGAISIVGKKLSEVNYILQSR  
TQFIELDKKGIGIRQMSALKERAGILITDQDCVENESVRLKFNKENVYOMGASL  
IAQRNALIIPYATGEGKFCIEFFKAKDSQNASLEELTYQAQSCSEMIKRPWEY  
FFHRRFASYNEEIYKAK"

1222..2769

/note="orf3"

/codon\_start=1

/transl\_table=11

/product="putative two-domain glycosyltransferase"

/protein\_id="AAL05986.1"

/db\_xref="GI:15718485"

/translation="WNLKQISVIIIVKNAEQTLLECLNSLKDQFDEIILLNESIDNTL  
NIANFKDFANLYVHNAFIGFAGKLNLSVAKNDWILSIDADEVLENECIKELKN  
LELQDNLIILSRKNLYKGEWIKACGWSPDYVLRIPKNFTRFNDNLVHESLVLPSNV  
KLKHNGYKHYAFRDI SHLKNQYISLWAKQNHKKSGVLKANURAFWTFRNTF  
LKNGLYGYQGCTISVCNALGTFKYMKLQKQKPKCALITTYNQERLKLVID  
SVKNLTFLPNEVLIIADGSKEDTARLIEYQKDFPGCSLKHWOEDGFRAAKRNNAI  
RVSKSEYIILIDGDMILEKDFIKNHFEFAKRVKLQGSRVILNKNESEEILKDDYSL

AFSKDPKSKNSFLAKIYKISKIDQKFNTEIILKSGSKTCNMSFFKADFBELDGN  
ENFIGNGDESEFVARFLNKGIFRRLKPKAIAYIHYESKNMLESNHOIYLDTIK  
NKKISWR"

CDS

2760..3944  
/note="orf4"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAL05987.1"  
/db\_xref="GI:15718486"

/translations="MEIKHEENRCVYNIORRIFERMENSVINOTVNLIELVNDG  
STDESLNAKEYTLKDERFLDKNGGLSSARNVGEYFSGEYKLNKTHIKENS  
LIEFQLDGNPNYIKAYKSSQAPNEKDLTNFTYNIIDYIIFLSDSDYELNCEIEC  
ISGRKVDVLFWDHDCYEDNIKNKHKTRMEIFDFKECIIIPKSYANRALSVGSRD  
VFGWMDIDFNFKIKLKFIPNFIINEDIHFGIILFASANKYIILSKYLCLRLRAN  
SISNHDKITKANVSFYKDIYETFGENAKEAKNYLKAASRVTTALKLIEFFKDOKE  
NALAIKETPLPCVAKKALMIKKFKDPLNLKEQLVLKPIQTKLPVDIWKFWQIKK  
IN"

gene

3988..4248  
/gene="cgtA-I"  
3988..4248  
/gene="cgtA-I"

CDS

/note="orf5-I; premature translation stop due to a one bp  
deletion"

/codon\_start=1  
/transl\_table=11  
/product="truncated  
beta-1,4-N-acetylglactosaminyltransferase"  
/protein\_id="AAL05988.1"  
/db\_xref="GI:15718487"

/translations="MLFQSFVKIICLFPFRKIRHKIKTFLLNKIQRDKIDSYLPK  
KILIQKYNNEDLINKLKAITGGIKDILIMMKNLKIQLL"

gene

/gene="cgtB"  
complement(5016..5921)  
complement(5016..5921)  
/gene="cgtB"

CDS

/codon\_start=1  
/transl\_table=11  
/product="beta-1,3-galactosyltransferase"  
/protein\_id="AAL05989.1"  
/db\_xref="GI:15718488"

/translations="MKISILPTYNVEQVIARAIESCINQTFKDIEIIVDDCGNDN  
SINIAKEYSKDKRIKIHNEKNLGLLRAGEYGVKANSPYIMFLDPDDYLELNACEE  
CKILDEQEDVIFVNAIVESNVIYKRPDFNSGYFSKFEVKYIAKNLYWTWVG  
KLIRKKLYEASLEKDKVKNMAEDVLLYPMLSQAKIAYMNCNLYHYVPNNNS  
ICWTNEVLVKNNIQELQVLNLYRONYLKNCYSVLVLYKLYLLYIYIKIRKTKLM  
VTLAKINILTLKILFKYKFLKQC"

gene

6010..6885  
/gene="cst-II"  
6010..6885  
/gene="cst-II"

CDS

/codon\_start=1  
/transl\_table=11  
/product="alpha-2,3-sialyltransferase"  
/protein\_id="AAL05990.1"  
/db\_xref="GI:15718489"

/translations="MKKVIIAGNPSLKEIDYSRLPNDPVRNCFYFEDKYVLGKK  
CTVFYTPNFFEQVYTLKHLQNOEYETELMCSYNQAHLENFVKTFYDFPDA  
HLGYDFKQKFNAPFKHEIYFNORIISGVYMCVAIALGYKEIYLSGIDFYQGS  
SYAFDTPQENLKLAPDFKNDRSHYIGHSKNTDIDKALEFLEKTYKILYCLCPNSLLA  
NFIELAPNLSNFIIOEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENNVYKLIKD  
LRLPSPDIKHYFKGK"

gene

6886..7926  
/gene="neuB1"  
6886..7926  
/gene="neuB1"

CDS

/codon\_start=1  
/transl\_table=11  
/product="sialic acid synthase"  
/protein\_id="AAL05991.1"  
/db\_xref="GI:15718490"

/translations="MKKEIKQNIIESEKAPLVVPEIGINHNSLELAKIMVDAFSA  
GAKLIKQTHIVEDEMSKAARKVIPGNAKISIVEIMQKCALDYKDELALKEYTEKLG

gene

CDS

VLTSPFSRAGANRLDMGVSAFKITGSGECNNYPLIKHIAAFKKPMIVSTGMNSIESI  
KPTAKILLDNEIPFVMTTNLYPTPHNLRLNMLELKEKFSQWGLSDHDTNLAC  
LGAVALGCVLERHPTDSMHRSGPDIVCSMDTKALKELIIQSEOMAIKMGNSKAA  
KOEQVTTIDFAPASVVSIKDIKKGEVLSDMDNIWVRPGLGGISAABENILGKKALRDI  
ENDTQLSYEDFA"

7923..9047  
/gene="neuCl"  
7923..9047  
/gene="neuCl"  
/codon\_start=1  
/transl\_table=11  
/product="putative N-acetylglucosamine-6-phosphate  
2-epimerase"  
/protein\_id="AAL05992.1"  
/db\_xref="GI:15718491"

gene

CDS

/translations="MKKILFITGTADYSKISKLVVRVONSSEFELYIPATGMHLSKN  
EGYTVKELYKNGFKNIYEFINYDKVYQTDKALATTIDGFSRYVNELKPLDIVHGBRI  
EPLAAIVGALNNILVAHIEGESITDIDSLRHAIKLAHILVNDFAKRLMQIG  
EDEKSIIFIIGSPDLELNDNKISLEAKKYDINYENALLMFPHPVTTITSIKQAD  
NLKALIQSNKNYIYIPNNDLGFELIQSYEELKNRPFKLFPSLRFFYFTLLKNA  
DFIIGNSCILKEALYLKTAGILVGRONGRLGNTLKVANSDEILKAINITIHKKQ  
DLFTKLEILDDSKLFPEYLLSGEFFKLSTQKVKDIKHA"

9040..9987  
/gene="cgtA-II"  
9040..9987  
/gene="cgtA-II"  
/note="orf5-II"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,4-N-acetylglactosaminyltransferase"  
/protein\_id="AAL05993.1"  
/db\_xref="GI:15718492"

/translations="MLKKIISLYKRYSTSKLVLONEHFIKENKNIYKGHKGFDFD  
EKADKSPNLPMGFIRVKNALTLRVSLSTLPALORGIAYNDCDDGSEELILEFC  
KQYENFIKKYPYKVDLENPKNEENKLSYNNWAASFPLDEWFIKIDVDHYDAKLL  
YKSPYRIDQENKALCYPRINFILNGNIYVQNSNGYFGIGGQDLILKRNSSFIERR  
VSKSOWIDPKLJIEEYSEQQVLSQGVLIQAPLLOWHFPALPYRRNDYQQYLDILS  
LEEFQAFHRKSKEARKIDFAMLRPVIQILKKFQGEIK"

gene

CDS

9984..10649  
/gene="neuA1"  
9984..10649  
/gene="neuA1"  
/codon\_start=1  
/transl\_table=11

Query Match 100.0%; Score 873; DB 15; Length 12390;  
Best Local Similarity 100.0%; Pred. No. 6.6e-215; Mismatches 0; Indels 0; Gaps 0;  
Matches 873; Conservative 0;

Qy 1 ATGAAAAAAGTTATTATTCTGGAATGGAACCAAGTTTAAAAAGAAATGATTATTCAAGG 60

Db 6010 ATGAAAAAAGTTATTATTCTGGAATGGAACCAAGTTTAAAAAGAAATGATTATTCAAGG 6069

Qy 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

Db 6070 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6129

Qy 121 CTGTGTAATAAATGCAAAACAGTGTTTACACCCCTAATTTCTTTCTTGAGCAATACTAC 180

Db 6130 CTGTGTAATAAATGCAAAACAGTGTTTACACCCCTAATTTCTTTCTTGAGCAATACTAC 6189

Qy 181 ACTTTAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAACTAATATGTGTTCTAAT 240

Db 6190 ACTTTAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAACTAATATGTGTTCTAAT 6249

Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACACTTTTACGATTTATTTTCCT 300

Db 6250 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACACTTTTACGATTTATTTTCCT 6309

Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTATTAAA 360

Db 6310 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTATTAAA 6369

```
QY 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
DB |||||||
QY 6370 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 6429
DB |||||||
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 480
DB |||||||
QY 6430 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGGGTCA 6489
DB |||||||
QY 481 TCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
DB |||||||
QY 6490 TCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6549
DB |||||||
QY 541 GATCCTCACACTATATCGGCATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 600
DB |||||||
QY 6550 GATCCTCACACTATATCGGCATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 6609
DB |||||||
QY 601 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTAAATGCTTTTAGCAATTTT 660
DB |||||||
QY 6610 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTAAATGCTTTTAGCAATTTT 6669
DB |||||||
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATAACAAGAAAAATAACTACACT 720
DB |||||||
QY 6670 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATAACAAGAAAAATAACTACACT 6729
DB |||||||
QY 721 AAAGATATACATACCTTCTAGTGAGGCTTATGCAAAATTTTCAAAAAATATTAATTT 780
DB |||||||
QY 6730 AAAGATATACATACCTTCTAGTGAGGCTTATGCAAAATTTTCAAAAAATATTAATTT 6789
DB |||||||
QY 781 AAAAAATATAAAATTTAAAGAAATCTTTTATACAAAGTTGATATAAGATCTATTAAAGATTA 840
DB |||||||
QY 6790 AAAAAATATAAAATTTAAAGAAATCTTTTATACAAAGTTGATATAAGATCTATTAAAGATTA 6849
DB |||||||
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
DB |||||||
QY 6850 CCTAGTGATATAAGCAATTTTCAAGGAAAA 6882
DB |||||||

RESULT 8
AF305571 4749 bp DNA linear BCT 29-JAN-2002
LOCUS Campylobacter jejuni strain 81-176 sialyl transferase (cst), sialic
DEFINITION acid synthase (neuB), NeuC1, N-acetylglucosyltransferase, and
CMF-NeuNac synthetase (neuA) genes, complete cds.
ACCESSION AF305571
VERSION AF305571.1 GI:15824490
KEYWORDS Campylobacter
SOURCE Campylobacter jejuni subsp. jejuni 81-176
ORGANISM Campylobacter jejuni subsp. jejuni 81-176
REFERENCE Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
AUTHORS Campylobacteraceae; Campylobacter.
1 (bases 1 to 4749)
Guerry, P., Symanski, C.M., Prendergast, M.M., Hickey, T.E.,
Ewing, C.P., Pattarini, D.L. and Moran, A.P.
TITLE Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide
affects ganglioside mimicry and invasiveness in vitro
JOURNAL Infect. Immun. 70 (2), 787-793 (2002)
PUBMED 11796612
REFERENCE 2 (bases 1 to 4749)
Guerry, P., Symanski, C.M., Prendergast, M.M., Hickey, T.E.,
Ewing, C.P., Pattarini, D.L. and Moran, A.P.
AUTHORS Direct Submission
TITLE Submitted (14-SEP-2000) Enteric Diseases, Naval Medical Research
JOURNAL Center, 503 Robert Grant Ave, Silver Spring, MD 20910, USA
FEATURES Location/Qualifiers
source 1..4749
/organism="Campylobacter jejuni subsp. jejuni 81-176"
/mol_type="genomic DNA"
/strain="81-176"
/db_xref="taxon:354242"
89..964
/gene="cst"
89..964
/gene="cst"

gene
CDS
```

ORIGIN

```
/note="Cst"
/codon_start=1
/transl_table=11
/product="sialyl transferase"
/protein_id="AAL09368.1"
/db_xref="GI:15824491"
/translation="MKKVIAGNGPSLKEIDYSLPNDFDVFRCNQFVFEKYLGLKK
CRAVFYTPFFPEQYYTLKHLIQEYETELIMCSNTYNQAHLENENFVKTFDYFPDA
HGUDYFKQLKEFNAYKFEHIFYNQIRTSQVMCAVAIALGYKEIYLSGDDFYONGS
SYAFDPKQNLKLAPDKFNDKRSYIGHKSHYIKALEFLEKTKYIKLYCLCPNSLLA
NFTELAPNLSNFIIQEKNNYTKDILIPSEAYGKFSKNINFKIKIKENVVYKLIKD
LRLPSDIKHYPKGK"
965..2005
/gene="neuB1"
965..2005
/gene="neuB1"
/notes="NeuB1"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAL09369.1"
/db_xref="GI:15824492"
/translation="MKEIKQNIISSEKAPLVVPEIGINHNGSLLELAKIMVDAFSA
GAKIKHQTHIVEDENSKAAKVIPOGNAKISYIEIMOKCALDYKDELAKYTEKGLG
VILSTPFRAGANRLSDMGVSAFKIGSGECNNYPLIKHIAAFKPKMIVSTGWNSTESI
KPTVKILLNEIPFVLMHSTNLPTPHNLVRLNMLKKEFEKFCMVGSLSDRTDNLAC
LGAVALGACVLRHFTHDSMRSGPDIVCSMDTKALKELIIQSEQMAIMKGNESKAA
KOBQVTTIDFAPASVWSIKDIKKGVELSMDNIWVKRPGLGISAAFEFENILGCKALRDI
ENDTQLSYEDFA"
2062..3126
/notes="N-acetylmannosamine synthesis"
/codon_start=1
/transl_table=11
/product="NeuC1"
/protein_id="AAL09370.1"
/db_xref="GI:15824493"
/translation="MYRVQNSSEFELYIFATGMHLSKNFGYTVKELYKNGKPNIEFEI
NYDKYQTDKALATTIDGFSRYVNEKLDPDIIVHGDRIEPLAAIIVGALNNILVAHIE
GGEISQTDLSRHAISKLAIHLVNDFAKRLMOLGEDEKSIFFIGSPDLELNDN
KISLSEAKKYDINVENYALLMEHPVTTETSIKNOADNLVKALIQSNKNVIVLPNN
DIGLFSRLQSYELKNNPRKLPFSPRFEYITLLKNADFIIGNSCILKALFLKTA
GIVSGRQNGRQENLTKVANSDEILKAINTIHKKQDLFTSKLEILDSSKLFEYL
LSGEFPKJSTQKVFQDKHA"
3119..4066
/notes="CstA"
/codon_start=1
/transl_table=11
/product="N-acetylglucosyltransferase"
/protein_id="AAL09371.1"
/db_xref="GI:15824494"
/translation="MLKKIISLVKRYISIKKLVDNEHFTKENKNIYKXKHGFFDFD
EKAKOVKSLNPWFIRVKNREALTVLSLESILPALQORGLIAYNDCDGGSEELILEPC
KQFPFIADKYPYKVDLENPKNEENKLIYINWAASFIPUDENFIKIDVHRYSDAKKL
YFVRIDENKALCYPRINFILNGNIYVQNSGYFIGGGQDLILKRRNYSFIERR
VSKSQWIDPKGLIELYSEQVLSQVILQAPLQWHPFALKYRRENDYQQYLDILS
LEEFQAFHRKSKSEAKKIDFAMLRKRPVIEQILKKFQGEIK"
4063..4728
/gene="neuA1"
4063..4728
/gene="neuA1"
/notes="NeuA1"
/codon_start=1
/transl_table=11
/product="CMP-NeuNac synthetase"
/protein_id="AAL09372.1"
/db_xref="GI:15824495"
/translation="MSLAIIPARGSGKIGKNKLVLNNKPLIYVTTIKAAINAKSISK
VWVSSDSIELNVAKSONVDILKRPISLAQDDTTSDKVLHLKFKYKDYEDVDFLOPT
SLRFTNIHDKAFNLYKSNANALISVSCDNKILKAFVNDYGLAGI CNDEYFPEP
ROKLPKTTMSNGAIYILKIFELNPFQNKTHFMDSESSDIDCLDELDKKVEQI
WKK"
```

Query Match		99.8%;	Score 871.4;	DB 15;	Length 4749;
Best Local Similarity		99.9%;	Pred. No. 2.7e-214;		
Matches 872;		Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	ATGAAAAAGCTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGG	60		
Db	89	ATGAAAAAGCTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGG	148		
QY	61	CTACCAATGATTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120		
Db	149	CTACCAATGATTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	208		
QY	121	CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTCTTTTGAGCAATCTAC	180		
Db	209	CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTCTTTTGAGCAATCTAC	268		
QY	181	ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCCGAACCTAAATATGTGTTCTAAT	240		
Db	269	ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCCGAACCTAAATATGTGTTCTAAT	328		
QY	241	TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT	300		
Db	329	TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT	388		
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAAGCTTATTTTAA	360		
Db	389	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAAGCTTATTTTAA	448		
QY	361	TTTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC	420		
Db	449	TTTTCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGAGTAGCC	508		
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA	480		
Db	509	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA	568		
QY	481	TCATTGCTTTTGATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	540		
Db	569	TCATTGCTTTTGATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT	628		
QY	541	GATCGCTCACATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA	600		
Db	629	GATCGCTCACATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA	688		
QY	601	GAATAAACTTACAAAAATAAACTATATTGCTTATGCTCTAATAGTCTTTTACGCAATTTT	660		
Db	689	GAATAAACTTACAAAAATAAACTATATTGCTTATGCTCTAATAGTCTTTTACGCAATTTT	748		
QY	661	ATAGAACTAGCGCAAAATTTAAATTTCAAATTTTATCATACAAGAAAAATAACTACA	720		
Db	749	ATAGAACTAGCGCAAAATTTAAATTTCAAATTTTATCATACAAGAAAAATAACTACA	808		
QY	721	AAAGATATATCTATCTCTAGTGGCTTATGGAATAATTTTCAAAAAATATTAATTT	780		
Db	809	AAAGATATATCTATCTCTAGTGGCTTATGGAATAATTTTCAAAAAATATTAATTT	868		
QY	781	AAAAAATAAATAAAGAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA	840		
Db	869	AAAAAATAAATAAAGAAATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA	928		
QY	841	CCTAGTGATATAAGCATTATTTTCAAGGAAAA	873		
Db	929	CCTAGTGATATAAGCATTATTTTCAAGGAAAA	961		

RESULT 9  
LOCUS AX934431  
DEFINITION Sequence 8 from Patent WO02074942.  
ACCESSION AX934431  
VERSION AX934431.1 GI:40641678  
KEYWORDS Campylobacter jejuni

ORGANISM	Campylobacter jejuni			
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.			
AUTHORS	1 Gilbert, M. and Wakarchuk, W.W.			
TITLE	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics			
JOURNAL	Patent: WO 02074942-A 8 26-SEP-2002; National Research Council of Canada (CA)			
FEATURES	Location/Qualifiers			
source	1. 876			
	/organism="Campylobacter jejuni"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:197"			
CDS	1. 876			
	/note="unnamed protein product; Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C. jejuni O:19"			
	/codon_start=1			
	/transl_table=11			
	/protein_id="CAF04169.1"			
	/db_xref="GI:40641679"			
	/translation="MKVIIAGNPSLKEIDYSRLPNDFVFCNQYFEDKYLKCKAVYTNFFPEQYITLHLIQNOEYETELIMCSNNOAHLENENFVKTYFYDFPDAHLGDDFFKQLKEFNAYFKHEIYFNORITSGVMCAVAIALGYKEIYLSIGDIFYQGS SYADFTQENLLKLAPDFKNDRSHYIGHSKNTDIKALELEKTKYKIKLYCLCPNSLLA NFIELAPNLNFIIOEKNNVTKDILIPSEAYGKFSKNINPKKIKENVYVKLIKDLLRLPSDIKHYFKGK"			
ORIGIN				
	Query Match	99.6%;	Score 869.8;	DB 2; Length 876;
	Best Local Similarity	99.8%;	Pred. No. 1.5e-213;	
	Matches 871;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
QY	1	ATGAAAAAGCTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGG	60	
Db	1	ATGAAAAAGCTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTCAAGG	60	
QY	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120	
Db	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT	120	
QY	121	CTTGTTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC	180	
Db	121	CTTGTTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTAC	180	
QY	181	ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATGTTGTTCTAAT	240	
Db	181	ACTTTAAACATTTAATCCAAATCAAGAAATAGAGACCGAACTAATATGTTGTTCTAAT	240	
QY	241	TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT	300	
Db	241	TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTTCCT	300	
QY	301	GATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	360	
Db	301	GATGCTCATTTGGGATATGATTTTAAAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	360	
QY	361	TTTCAGCAAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGTCAGTAGCC	420	
Db	361	TTTCAGCAAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGTCAGTAGCC	420	
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA	480	
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA	480	
QY	481	TCATTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540	
Db	481	TCATTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540	
QY	541	GATCGCTCACATATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA	600	
Db	541	GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA	600	



```

QY 601 GAAAAAAGTACAAATATAATGCTTATGCTTAATAGTCTTTTACCAATTTT 660
Db 601 GAAAAAAGTACAAATATAATGCTTATGCTTAATAGTCTTTTACCAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAAGAAAAAATAACTACACT 720
QY 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780
Db 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780
QY 781 AAAAAAATAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAGAATTA 840
Db 781 AAAAAAATAAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAGAATTA 840
QY 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873

RESULT 10
BD249794
LOCUS
DEFINITION
Campylobacter glycosyltransferases for biosynthesis of
gangliosides and ganglioside mimics.
ACCESSION
BD249794
VERSION
BD249794.1 GI:33059564
KEYWORDS
JP 2002535992-A/5.
SOURCE
Campylobacter jejuni
ORGANISM
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 876)
Gilbert,M. and Wakarchuk,W.W.
Campylobacter glycosyltransferases for biosynthesis of
gangliosides and ganglioside mimics
Patent: JP 2002535992-A 5 29-OCT-2002;
NATIONAL RESEARCH COUNCIL OF CANADA
OS
Campylobacter jejuni
PN
JP 2002535992-A/5
PD
29-OCT-2002
PF
01-FEB-2000 JP 2000597438
PR
01-FEB-1999 US 60/118213,31-JAN-2000 US 06/495406 PI
MICHEL GILBERT,WARREN W WAKARCHUK
PC
C12N15/09,C12N1/21,C12N9/10,C12N9/88,C12N15/00 CC
Campylobacter glycosyltransferases for biosynthesis of
gangliosides and
CC
ganglioside mimics
FH
Key
FT
source
Location/Qualifiers
1..876
/organism='Campylobacter jejuni'
FEATURES
source
Location/Qualifiers
1..876
/organism='Campylobacter jejuni'
/mol_type='genomic DNA'
/db_xref='taxon:197'
ORIGIN
Query Match 99.6%; Score 869.8; DB 2; Length 876;
Best Local Similarity 99.8%; Pred. No. 1.5e-213;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAAAAAGTTATATTGCTGGAATGGACCAAGTTTAAAGAAATGATATTCAAGG 60
Db 1 ATGAAAAAGTTATATTGCTGGAATGGACCAAGTTTAAAGAAATGATATTCAAGG 60
QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAGATAAATACTAT 120
QY 121 CTTGGTAAAAATGCAAAACAGTCTTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180

```

```

Db 121 CTTGGTAAAAATGCAAAACAGTCTTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180
QY 181 ACTTTAAAAATTAATCCAAAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTAAT 240
Db 181 ACTTTAAAAATTAATCCAAAAATCAAGAAATATGAGACCGAACTAAATATGTGTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTTCTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTTTCTCT 300
QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
QY 361 TTTTACGAAATTTATTTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Db 361 TTTTACGAAATTTATTTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGACATAGTAAAGAAATACAGATATAAGCTTTTAGAATTTCTA 600
Db 541 GATCGCTCGACTATATCGACATAGTAAAGAAATACAGATATAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAAGAAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAAGAAAAAATAACTACACT 720
QY 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780
Db 721 AAAGATATACATACCTTCTAGTGGGCTTATGGAATTTTCAAAAATTTTAAATTTT 780
QY 781 AAAAAAATAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAGATCTATTAGAATTA 840
Db 781 AAAAAAATAAAAATTAAGAAAAATGTTTATCAAGTTGATAAAGATCTATTAGAATTA 840
QY 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873

```

## RESULT 11

```

AR271703
LOCUS
DEFINITION
Sequence 8 from patent US 6503744.
ACCESSION
AR271703
VERSION
AR271703.1 GI:29703248
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 876)
AUTHORS
Gilbert,M. and Wakarchuk,W.W.
TITLE
Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL
Patent: US 6503744-A 8 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source
Location/Qualifiers
1..876
/organism='unknown'
/mol_type='genomic DNA'

```



Db	541	GATCGCTCGCACTATATCGGACACTAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA	600
Qy	601	GAAGAAACCTTACAAAATATAAATCTATATGCTTATGCTCTCTAATAGTCTTTTAGCAATTTT	660
Db	601	GAAGAAACCTTACAAAATATAAATCTATATGCTTATGCTCTAATAGTCTTTTAGCAATTTT	660
Qy	661	ATAGAACTAGCGCCAAATTTAAATTTCAAATTTTCAATTTTATCATACAAGAAAAAATAACTACACT	720
Db	661	ATAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATACAAGAAAAAATAACTACACT	720
Qy	721	AAAGATATATCTCATACCTCTAGTAGGCTTATCGAAAAATTTTCAAAAAATATTAATTTT	780
Db	721	AAAGATATATCTCATACCTCTAGTAGGCTTATCGAAAAATTTTCAAAAAATATTAATTTT	780
Qy	781	AAAAAATATAAAATTTAAAGAAAATGTTTATTACAAGTTGATAAAGATCTATTAGATT	840
Db	781	AAAAAATATAAAATTTAAAGAAAATGTTTATTACAAGTTGATAAAGATCTATTAGATT	840
Qy	841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
Db	841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
RESULT 14			
AR609663			
LOCUS	AR609663	876 bp	DNA linear PAT 15-DEC-2004
DEFINITION	Sequence 8 from patent US 6825019.		
ACCESSION	AR609663		
VERSION	AR609663.1	GI:56664963	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 876)		
AUTHORS	Gilbert, M. and Wakarchuk, W.W.		
TITLE	Polypeptides having beta-1,3-galactosyl transferase activity		
JOURNAL	Patent: US 6825019-A 8 30-NOV-2004;		
	National Research Council of Canada; Ottawa;		
	CAN;		
FEATURES	Location/Qualifiers		
source	1..876		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	99.6%;	Score 869.8;	DB 2; Length 876;
Best Local Similarity	99.8%;	Pred. No. 1.5e-213;	
Matches 871;	Conservative 0;	Mismatches 213;	Indels 0; Gaps 0;
Qy	1	ATGAAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAAAATTTGATTATTCAAGG	60
Db	1	ATGAAAAAAGTTATTATTCTGCGAAATGGACCAAGTTTAAAGAAAAATTTGATTATTCAAGG	60
Qy	61	CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT	120
Db	61	CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTGTAAGATAAATACTAT	120
Qy	121	CTTGTTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTCTTTTGACCAATACTAC	180
Db	121	CTTGTTAAAAATGCAAAACAGTGTTTTACACCCCTAATTTCTCTTTTGACCAATACTAC	180
Qy	181	ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAACTAAATATGTGTTCTAAT	240
Db	181	ACTTTAAAAACATTTAATCCAAATCAAGATATGAGACCGAACTAAATATGTGTTCTAAT	240
Qy	241	TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCCCT	300
Db	241	TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTTCCCT	300
Qy	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTAAGAAATTTAATGCTTATTTTAAA	360
Db	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACCTTTAAAGAAATTTAATGCTTATTTTAAA	360

```
QY 361 TTTCCAGAAATTTATTTCAATCAAGAAATACCTCAGGGCTCTATATGTGTGAGTAGCC 420
Db 361 TTTCCAGAAATTTATTTCAATCAAGAAATACCTCAGGGCTCTATATGTGTGAGTAGCC 420
QY 421 ATAGCCCTAGGATCAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATCAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGACATAGTAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGACATATCGACATAGTAAATACAGATATAAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTACAAATATAAATATATTTGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 660
Db 601 GAAAAAATTTACAAATATAAATATATTTGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACT 720
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACT 720
QY 721 AAGATATATCTATCTCTAGTGGCTTATGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 780
Db 721 AAGATATATCTATCTCTAGTGGCTTATGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 780
QY 781 AAAAAAATAAATAAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAATAAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTTTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAAA 873

RESULT 15
AR689941
LOCUS AR689941
DEFINITION Sequence 8 from patent US 6905867.
ACCESSION AR689941
VERSION AR689941.1 GI:74471949
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl
transferase activity
JOURNAL Patent: US 6905867-A 8 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 99.8%; Score 869.8; DB 2; Length 876;
Best Local Similarity 99.8%; Pred. No. 1.5e-213;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTGCTGGAATGGCCAAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTGCTGGAATGGCCAAAGTTTAAAGAAATTTGATTTTCAAGG 60
QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
QY 121 CTTGGTAAAAAATGCAAAAACAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
```

```
Db 121 CTTGGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
QY 181 ACTTTAAACAATTTAAATCCAAAATATAGAGCCGAACTAAATATATGTGTTCTAAT 240
Db 181 ACTTTAAACAATTTAAATCCAAAATATAGAGCCGAACTAAATATATGTGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTCC 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTATAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTTATTATAA 360
QY 361 TTTTCAACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATCAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATCAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCGACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTACAAATAAATAAATACTATATTTGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 660
Db 601 GAAAAAATTTACAAATAAATAAATACTATATTTGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 660
QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACT 720
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACT 720
QY 721 AAGATATATCTATCTCTAGTGGCTTATGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 780
Db 721 AAGATATATCTATCTCTAGTGGCTTATGCTTATGTCTCTAATAGTCTTTTAGCAATTTT 780
QY 781 AAAAAAATAAATAAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAAATAAATAAAGAAATGTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTTTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAAA 873
```

Search completed: July 19, 2006, 19:12:28  
Job time : 5053.55 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:53:32 ; Search time 652.919 Seconds  
(without alignments)  
9322.405 Million cell updates/sec

Title: US-10-734-719-13

Perfect score: 873

Sequence: 1 atgaaaaaagtattattgc.....agcattatttcaaggaaaa 873

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 348612431 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N\_Geneseq\_8.\*  
1: Geneseqnl980s.\*  
2: Geneseqnl990s.\*  
3: Geneseqnl2000s.\*  
4: Geneseqnl2001as.\*  
5: Geneseqnl2001bs.\*  
6: Geneseqnl2002as.\*  
7: Geneseqnl2002bs.\*  
8: Geneseqnl2003as.\*  
9: Geneseqnl2003bs.\*  
10: Geneseqnl2003cs.\*  
11: Geneseqnl2003ds.\*  
12: Geneseqnl2004as.\*  
13: Geneseqnl2004bs.\*  
14: Geneseqnl2005s.\*  
15: Geneseqnl2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	873	100.0	873	8	ABT13671 C. jejuni
2	869.8	99.6	876	3	AAAS3726 Campyloba
3	869.8	99.6	876	8	ABT13669 C. jejuni
4	866.6	99.3	873	8	ABT13670 C. jejuni
5	841	96.3	876	3	AAAS3724 Campyloba
6	841	96.3	876	8	ABT13667 C. jejuni
7	839.4	96.2	876	13	ADU77566 Sialytran
8	839.4	96.2	12219	13	ADU77595 Campyloba
9	834.6	95.6	11474	3	AAAS3720 LPS core
10	834.6	95.6	11474	8	ABT13665 Campyloba
11	833	95.4	876	3	AAAS3721 Campyloba
12	833	95.4	876	8	ABT13666 C. jejuni
13	817	93.6	876	3	AAAS3725 Campyloba
14	817	93.6	876	8	ABT13668 C. jejuni
15	311.6	35.7	1293	2	AAZ25693 Campyloba
C 16	286.8	32.9	110000	2	Continuation (4 of
C 17	282.6	32.4	191996	13	ADT05647 Haemophil
C 18	265.6	30.4	13379	13	ADT05530 Haemophil

C 19	265.6	30.4	117576	13	ADT05736 Haemophil
C 20	265.6	30.4	349980	13	ADT05648 Haemophil
C 21	262	30.0	4277	13	ADT05421 Haemophil
C 22	255	29.2	909	8	ACC71694 alpha-2,3
C 23	89.4	10.2	8056	8	ABZ10246 Haematopo
C 24	89.2	10.2	8056	8	ABZ10246 Haematopo
C 25	87	10.0	2482	14	ADZ71091 Human chr
C 26	86.2	9.9	19087	6	ABL32793 Human imm
C 27	85.2	9.8	8056	8	ABZ10100 Haematopo
C 28	82.6	9.5	158001	12	ADL17884 Human pho
C 29	81.2	9.3	1554	14	ADZ70999 Human chr
C 30	80.6	9.2	15548	6	ABL34155 Human imm
C 31	79.6	9.1	8056	8	ABZ10100 Haematopo
C 32	78.2	9.0	6465	8	ABL32985 Human imm
C 33	77.2	8.8	1501	8	ABZ10188 Haematopo
C 34	77.2	8.8	1501	10	ADZ84162 Human lym
C 35	76.8	8.8	17131	6	ABL33053 Human lym
C 36	75.6	8.7	18988	4	AAAS46342 Tumour su
C 37	75.6	8.7	18988	6	ABL32701 Human imm
C 38	75.6	8.7	18988	6	ABL34509 Human'met
C 39	75.6	8.7	18988	6	ABL70204 Chemical
C 40	75.6	8.7	18988	7	ADS99770 Complement
C 41	75.4	8.6	5930	6	ABL32517 Human imm
C 42	74.2	8.5	19380	6	AAAS61427 Human gen
C 43	74	8.5	5979	4	AAAS45313 Chemical
C 44	74	8.5	5979	6	ADK28152 DNA trans
C 45	74	8.5	6132	6	ABL32863 Human imm

## ALIGNMENTS

### RESULT 1

ABT13671

ID ABT13671 standard; DNA; 873 BP.

XX ABT13671;

DT 07-FEB-2003 (first entry)

DE C. jejuni bifunctional sialtransferase cstII coding sequence #6.

KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.

XX Campylobacter jejuni.

XX WO200274942-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-CA0000229.

XX 21-MAR-2001; 2001US-00816028.

XX (CANA ) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2003-040554/03.

XX P-PSDB; ABU18485.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimetics.

XX Disclosure; Page 99; 107pp; English.

XX The invention comprises the amino acid and coding sequences of

CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acetyltransferase; glycosyltransferase; GalNAC (N-  
CC Acetyl-galactosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthetase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
CC gene of the invention  
XX  
SQ Sequence 873 BP; 353 A; 117 C; 107 G; 296 T; 0 U; 0 Other;  
Query Match 100.0%; Score 873; DB 8; Length 873;  
Best Local Similarity 100.0%; Pred. No. 3.2e-133; Indels 0; Gaps 0;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGCTTATTGCTGCAAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60  
DB 1 ATGAAAAAGCTTATTGCTGCAAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60  
QY 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
DB 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
QY 121 CTGGTAAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTTGGCAATACATC 180  
DB 121 CTGGTAAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTTGGCAATACATC 180  
QY 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240  
DB 181 ACTTTAAAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCC 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCC 300  
QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360  
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360  
QY 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATTTGATTTTATCAAAATGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTATACCAACCAAGAAATTTTAAACTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTATACCAACCAAGAAATTTTAAACTAGCCCTGATTTTAAAT 540  
QY 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600  
QY 601 GAAAAAATTTACAAAATAAAATTTATGCTTATGCTTCTTAATAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTTACAAAATAAAATTTATGCTTATGCTTCTTAATAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAGAAATAAATACTACACT 720  
DB 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAGAAATAAATACTACACT 720  
QY 721 AAAGATATCTCATCTCTAGTGGCTTATGGAAATTTTCAAAATATTTATTTT 780  
DB 721 AAAGATATCTCATCTCTAGTGGCTTATGGAAATTTTCAAAATATTTATTTT 780  
QY 781 AAAAAAATAAAATAAAGAAATTTTATTTATCAAGTTGATAAAGATCTTATTAAGATTA 840

DB 781 AAAAAAATAAAATAAAGAAATTTTATTTATCAAGTTGATAAAGATCTTATTAAGATTA 840  
QY 841 CCTAGTATATAAAGCATTTATTTCAAGGAAAA 873  
DB 841 CCTAGTATATAAAGCATTTATTTCAAGGAAAA 873  
RESULT 2  
ID AA53726 standard; DNA; 876 BP.  
XX AA53726;  
XX 15-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE Campylobacter jejuni O:19 serotype CstII sialyltransferase.  
XX  
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside; ds.  
XX  
XX Campylobacter jejuni; O:19 serotype.  
XX  
FH Location/Qualifiers  
FT CDS  
FT 1..876  
FT /\*tag= a  
FT /product= "CstII sialyltransferase"  
XX  
XX WO200046379-A1.  
XX  
XX 10-AUG-2000.  
XX  
XX 01-FEB-2000; 2000WO-CA000086.  
XX  
XX 01-FEB-1999; 99US-0118213P.  
XX 31-JAN-2000; 2000US-00495406.  
XX  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX  
XX WPI: 2000-524418/47.  
XX P-PSDB; AAY97212.  
XX  
XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
XX reagents and as immunogen for producing antibodies.  
XX  
XX Disclosure; Page 96; 120pp; English.  
XX  
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
XX useful for synthesising sialylated oligosaccharide such as ganglioside,  
XX lysoganglioside or their mimics. Glycosyltransferases are useful for  
XX chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
XX other oligosaccharides that have biological activity. The enzymes and  
XX nucleic acids that encode them are useful for studies of the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimics, such as C.  
XX jejuni and the nucleic acids are used as probes to study expression of  
XX genes involved in ganglioside mimetic synthesis. Antibodies raised  
XX against the glycosyltransferases are also useful for analyzing the  
XX expression patterns of these genes involved in pathogenesis. The nucleic  
XX acids are also useful for designing antisense oligonucleotides for  
XX inhibiting expression of the Campylobacter enzymes that are involved in  
XX the biosynthesis of ganglioside mimics that can mask the pathogens from  
XX the host's immune system. The oligosaccharides are useful as diagnosing  
XX reagents or as therapeutics and as immunogens for producing antibodies.  
XX Bacterial glycosyltransferase can be used to catalyse the formation of  
XX oligosaccharides that are identical to the corresponding mammalian  
XX structures and are easier and less expensive to produce in large

CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (Updated on 15-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Query Match 99.6%; Score 869.8; DB 3; Length 876;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-132;  
 Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 Db 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 Qy 61 CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 Db 61 CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 Qy 121 CTGTTGTAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 Db 121 CTGTTGTAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 Qy 181 ACTTTAAACATTTTAATCCAAATCAAGATATGAGCGGAATTAATGTTCTTAAT 240  
 Db 181 ACTTTAAACATTTTAATCCAAATCAAGATATGAGCGGAATTAATGTTCTTAAT 240  
 Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTGCT 300  
 Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTGCT 300  
 Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
 Db 301 GATGCTCATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
 Qy 361 TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGCTATATGTTGTCAGTACCC 420  
 Db 361 TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGCTATATGTTGTCAGTACCC 420  
 Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTC 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTC 480  
 Qy 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 Db 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 Qy 541 GATGCTCAGCTATATCGGACATAGTAAATACAGATATAAAGCTTTTGAATTTCTA 600  
 Db 541 GATGCTCAGCTATATCGGACATAGTAAATACAGATATAAAGCTTTTGAATTTCTA 600  
 Qy 601 GAAAAAATTTTCAAAATAAATACTATTTGCTTATGCTCTAATAGTCTTTTACGAAATTT 660  
 Db 601 GAAAAAATTTTCAAAATAAATACTATTTGCTTATGCTCTAATAGTCTTTTACGAAATTT 660  
 Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAAATCACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAAATCACT 720  
 Qy 721 AAAGATATCTATCACTTCTAGTGGCTTATGGAATAATTTTCAAAATAATTAATTT 780  
 Db 721 AAAGATATCTATCACTTCTAGTGGCTTATGGAATAATTTTCAAAATAATTAATTT 780  
 Qy 781 AAAAAATAAATAAATAAAGAAATTTTATTTACAGTTGATAAAGATCTATTAGATTA 840  
 Db 781 AAAAAATAAATAAATAAAGAAATTTTATTTACAGTTGATAAAGATCTATTAGATTA 840  
 Qy 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAA 873  
 Db 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAA 873

ABT13669

ID ABT13669 standard; DNA; 876 BP.

XX AC ABT13669;

XX DT 07-FEB-2003 (first entry)

XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #4.  
 XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNac transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.

XX PN WO200274942-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-CA0000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX PA (CANA ) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX DR WPI; 2003-040554/03.

XX DR P-PSDB; ABJ18482.

XX PT New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX PS Disclosure; Page 98; 107pp; English.

XX CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNac (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention

SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Query Match 99.6%; Score 869.8; DB 8; Length 876;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-132;  
 Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 Db 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
 Qy 61 CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 Db 61 CTACCAATGATTTTGATGTTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 Qy 121 CTGTTGTAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 Db 121 CTGTTGTAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTCTTTGAGCAATACTAC 180  
 Qy 181 ACTTTAAACATTTTAATCCAAATCAAGATATGAGCGGAATTAATGTTCTTAAT 240







Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAATACTACACT 720  
 QY 721 AAAGATATACCTATACCTCTCTAGTGGCTTTATGGAATAATTTTCAAAAAATTTAAATTTT 780  
 Db 721 AAAGATATACCTATACCTCTCTAGTGGCTTTATGGAATAATTTTCAAAAAATTTAAATTTT 780  
 QY 781 AAAAAATAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAGATTA 840  
 Db 781 AAAAAATAAAATTAAGAAAAATGTTTATACAAAGTTGATAAAGATCTATTAAGATTA 840  
 QY 841 CCTAGTGTATTAAGCAATTTTCAAAAGGAAA 873  
 Db 841 CCTAGTGTATTAAGCAATTTTCAAAAGGAAA 873

RESULT 6

ABT13667  
 ID ABT13667 standard; DNA; 876 BP.  
 XX

AC ABT13667;

DT 07-FEB-2003 (first entry)

DE C. jejuni bifunctional sialtransferase cstII coding sequence #2.

KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.

XX Campylobacter jejuni.

OS WO200274942-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-CA000229.

XX 21-MAR-2001; 2001US-00816028.

XX (CANADA) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2003-040554/03.

XX P-PSDB; ABJ18480.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.

PS Claim 8; Page 96-97; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention

SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

Query Match 96.3%; Score 841; DB 8; Length 876;

Best Local Similarity 97.7%; Pred. No. 5.2e-128;  
 Matches 853; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTCGATTATTCAGG 60  
 Db 1 ATGAAAAAGTTTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTCGATTATTCAGG 60  
 QY 61 CTACCAATGATTTTGTGATGTTTATAGATGTTTCAATTTTATTTTGAAGATAAATACTAT 120  
 Db 61 CTACCAATGATTTTGTGATGTTTATAGATGTTTCAATTTTATTTTGAAGATAAATACTAT 120  
 QY 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATACTAC 180  
 Db 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGGAGCAATACTAC 180  
 QY 181 ACTTTTAAACATTTTAAATCCAAATCAAGATATGAGCCGAACTAATTAATGTTCTTAAT 240  
 Db 181 ACTTTTAAACATTTTAAATCCAAATCAAGATATGAGCCGAACTAATTAATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCCT 300  
 Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCCT 300  
 QY 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTTAAAGAAATTTAAATGCTTATTTAAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAATTTTAAAGAAATTTAAATGCTTATTTAAA 360  
 QY 361 TTTCCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420  
 Db 361 TTTCCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCT 420  
 QY 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480  
 QY 481 TCTTATGCTTTTGTATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
 Db 481 TCTTATGCTTTTGTATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
 QY 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600  
 Db 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600  
 QY 601 GAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 660  
 Db 601 GAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 660  
 QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTCAATTTTATCATACAGAAAAATACTACACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTTAAATTTTCAATTTTATCATACAGAAAAATACTACACT 720  
 QY 721 AAAGATATACCTATCTCTAGTGGCTTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780  
 Db 721 AAAGATATACCTATCTCTAGTGGCTTTATGGAATAATTTTCAAAAAATTTTAAATTTT 780  
 QY 781 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 840  
 Db 781 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 840  
 QY 841 CCTAGTGTATTAAGCAATTTTCAAAAGGAAA 873  
 Db 841 CCTAGTGTATTAAGCAATTTTCAAAAGGAAA 873

RESULT 7

ADU77566  
 ID ADU77566 standard; DNA; 876 BP.  
 XX

AC ADU77566;

XX 10-FEB-2005 (first entry)

XX Sialyltransferase (cstII).

XX gene amplification; campylobacter infection; Sialytransferase; cstII; ds.  
XX Campylobacter jejuni.  
XX WO2004101817-A1.  
XX 25-NOV-2004.  
XX 19-MAY-2004; 2004WO-AU000654.  
XX 19-MAY-2003; 2003AU-00902445.  
XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
XX Shi F, Fry BN, Coloe PJ;  
XX WPI; 2004-834013/82.  
XX Identifying Campylobacter in a sample to comprises comparing the profile  
PT of an amplification product to a comparator Campylobacter species after  
PT performing nucleic acid amplification of the wla gene cluster.  
XX Claim 6; SEQ ID NO 13; 75pp; English.  
XX The invention describes a method of identifying Campylobacter in a sample  
CC comprising comparing the profile of an amplification product to a  
CC comparator Campylobacter species after performing nucleic acid  
CC amplification. Identifying Campylobacter in a sample comprises: (a)  
CC performing nucleic acid amplification by contacting the sample with a  
CC pair of nucleic acid primers, where at least one nucleic acid primer is  
CC derived from the wla gene cluster of Campylobacter, for a time and under  
CC conditions for generation of an amplification product comprising the wla  
CC gene cluster of Campylobacter or its portion; (b) preparing a profile of  
CC the amplification product; and (c) comparing the profile to a profile of  
CC a comparator Campylobacter species, strain or type, and thus, determining  
CC the identity of Campylobacter in the sample. Also described are: (a) a  
CC pair of nucleic acid primers, where at least one nucleic acid primer is  
CC derived from the wla gene cluster of Campylobacter; or (b) a pair of  
CC nucleic acid primers for amplification of the wla gene cluster of  
CC Campylobacter or its portion; or (c) a probe capable of hybridizing to  
CC the wla gene cluster of Campylobacter. The method and kit are useful for  
CC identifying and detecting the presence of Campylobacter in a sample, or  
CC for diagnosing infections associated with Campylobacter. This sequence  
CC represents Sialytransferase, encoded by a gene in the Campylobacter  
CC jejuni wla cluster used in the inventive method of detecting  
CC Campylobacter jejuni infection.  
XX  
SQ Sequence 876 BP; 356 A; 115 C; 104 G; 301 T; 0 U; 0 Other;  
Query Match 96.2%; Score 839.4; DB 13; Length 876;  
Best Local Similarity 97.6%; Pred No. 9.4e-128;  
Matches 852; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
DB 1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
QY 61 CTACCAATGATTGATGATTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120  
DB 61 CTACCAATGATTGATGATTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120  
QY 121 CTGTGTAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTCGACGAATCTAC 180  
DB 121 CTGTGTAATAATGCAAAACAGTGTGTTTACCAATCTCTGGTCTTTTTTTTGAACAACTACTAC 180  
QY 181 ACTTTAAACAATTTAATCAAAATCAAGAATATGACCGCACTAATATTGTTCTAAT 240  
DB 181 ACTTTAAACAATTTAATCAAAATCAAGAATATGACCGCACTAATATTGTTCTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAAA 360  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAAA 360  
QY 361 TTTTACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGACGTAGCC 420  
DB 361 TTTTACGAAATTTTATCTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGACGTAGCT 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGGAATTTGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
DB 481 TCTTATGCTTTTGATACAAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCACATATATCGGCACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600  
DB 541 GATCGCTCACATATATCGGCACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAACTTTACAAAAATAAAACTATATTGCTTATGTCCTAAATAGTCTTTTACAAATTTT 660  
DB 601 GAAAAAACTTTACAAAAATAAAACTATATTGCTTATGTCCTAAACAGTCTTTTACAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACACT 720  
DB 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACACT 720  
QY 721 AAGATATATCTATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780  
DB 721 AAGATATATCTATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780  
QY 781 AAAAAATAAAAAATTAAGAAATTTTATTTATACAGTTTGAATAAGATCTATTAAGATTA 840  
DB 781 AAAAAATAAAAAATTAAGAAATTTTATTTATACAGTTTGAATAAGATCTATTAAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
RESULT 8  
ADU77595  
ID ADU77595 standard; DNA; 12219 BP.  
AC ADU77595;  
XX  
DT 10-FEB-2005 (first entry)  
XX  
DE Campylobacter jejuni strain ATCC 43469 wla cluster.  
XX  
KW gene amplification; campylobacter infection; wla; ds.  
XX  
OS Campylobacter jejuni.  
XX  
PN WO2004101817-A1.  
XX  
PD 25-NOV-2004.  
XX  
PF 19-MAY-2004; 2004WO-AU000654.  
XX  
PR 19-MAY-2003; 2003AU-00902445.  
XX  
PA (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
XX  
PI Shi F, Fry BN, Coloe PJ;  
XX  
DR WPI; 2004-834013/82.  
XX  
PT Identifying Campylobacter in a sample to comprises comparing the profile  
PT of an amplification product to a comparator Campylobacter species after  
PT performing nucleic acid amplification of the wla gene cluster.

XX  
PS  
XX

Disclosure; Fig 6; 75pp; English.

The invention describes a method of identifying Campylobacter in a sample comprising comparing the profile of an amplification product to a comparator Campylobacter species after performing nucleic acid amplification. Identifying Campylobacter in a sample comprises: (a) performing nucleic acid amplification by contacting the sample with a pair of nucleic acid primers, where at least one nucleic acid primer is derived from the wla gene cluster of Campylobacter, for a time and under conditions for generation of an amplification product comprising the wla gene cluster of Campylobacter or its portion; (b) preparing a profile of the amplification product; and (c) comparing the profile to a profile of a comparator Campylobacter species, strain or type, and thus, determining the identity of Campylobacter in the sample. Also described are: (a) a pair of nucleic acid primers, where at least one nucleic acid primer is derived from the wla gene cluster of Campylobacter; or (b) a pair of nucleic acid primers for amplification of the wla gene cluster of Campylobacter or its portion; or (c) a probe capable of hybridizing to the wla gene cluster of Campylobacter. The method and kit are useful for identifying and detecting the presence of Campylobacter in a sample, or for diagnosing infections associated with Campylobacter. This sequence represents a Campylobacter jejuni strain ATCC 43469 wla cluster.

XX SQ Sequence 12219 BP; 4778 A; 1598 C; 1745 G; 4098 T; 0 U; 0 Other;

Query Match 96.2%; Score 839.4; DB 13; Length 12219;  
Best Local Similarity 97.6%; Pred. No. 9.2e-128;  
Matches 852; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
Dy 5995 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 6054  
Qy 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATACTAT 120  
Dy 6055 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATACTAT 6114  
Qy 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGGACAACTATAC 180  
Dy 6115 CTTGGTAAAAATTCAAAGCAGTATTTTACAACTCTGCTCTTTTGTGAACATACTAC 6174  
Qy 181 ACTTTAAACATTTAATTCCAAATCAAGAATATGAGACCGAAGTAATATGTTCTAAT 240  
Dy 6175 ACTTTAAACATTTAATTCCAAATCAAGAATATGAGACCGAAGTAATATGTTCTAAT 6234  
Qy 241 TACAAACCAAGCTCATCTAGAAATCAAAATTTTCTGAAAGTAAATTTTACGATTATTCCT 300  
Dy 6235 TACAAACCAAGCTCATCTAGAAATCAAAATTTTCTGAAAGTAAATTTTACGATTATTCCT 6294  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 360  
Dy 6295 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 6354  
Qy 361 TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATATGTCAGTAGCC 420  
Dy 6355 TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATATGTCAGTAGCT 6414  
Qy 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Dy 6415 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 6474  
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Dy 6475 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 6534  
Qy 541 GATGCTCACAATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600  
Dy 6535 GATGCTCACAATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 6594  
Qy 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAATAGCTTTTACGAAATTT 660  
Dy 6595 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAATAGCTTTTACGAAATTT 6654

Qy	661	ATAGAACTAGCGCCAAATTTAAATTTTATCATACAGAAAAATAACTACACT	720
Dy	6655	ATAGAACTAGCGCCAAATTTAAATTTTATCATACAGAAAAATAACTACACT	6714
Qy	721	AAAGATATACATACCTCTCTAGTGGCTTATGGAATTTTCAAAAAATATAATTTT	780
Dy	6715	AAAGATATACATACCTCTCTAGTGGCTTATGGAATTTTCAAAAAATATAATTTT	6774
Qy	781	AAAAAAATAAATTAAGAAAAATGTTTATTAACAAGTTCATAAAGATCTATTAAAGATTA	840
Dy	6775	AAAAAAATAAATTAAGAAAAATGTTTATTAACAAGTTCATAAAGATCTATTAAAGATTA	6834
Qy	841	CCTAGTGATATAAGCAATTTATTTCAAGGAAAA	873
Dy	6835	CCTAGTGATATAAGCAATTTATTTCAAGGAAAA	6867
RESULT 9			
AAA53720			
ID	AAA53720	standard; DNA; 11474 BP.	
XX	AC	AAA53720;	
XX	DT	15-SEP-2003 (revised)	
XX	DT	22-DEC-2000 (first entry)	
XX	DE	LPS core biosynthesis locus.	
XX	KW	Biosynthetic locus; biosynthesis; lipid A biosynthesis; acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase; Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase; sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody; immunity; immunogen; ganglioside; ds.	
XX	OS	Campylobacter jejuni; OH4384.	
XX	Key	Location/Qualifiers	
FT	CDS	350..1237	
FT		/tag= a	
FT		/product= "acyltransferase"	
FT		/note= "Open reading frame 2b"	
FT	CDS	1234..2490	
FT		/tag= b	
FT		/product= "glycosyl transferase"	
FT		/note= "Open reading frame 3a"	
FT	CDS	2786..3955	
FT		/tag= c	
FT		/product= "Glycosyl transferase"	
FT		/note= "Open reading frame 4b"	
FT	CDS	4025..5068	
FT		/tag= d	
FT		/product= "Beta 1,4-GalNAC transferase"	
FT		/note= "Open reading frame 5a"	
FT	CDS	5054..5959	
FT		/tag= e	
FT		/product= "Beta 1,2- galactosyltransferase"	
FT		/note= "Open reading frame 6a"	
FT	CDS	6048..6923	
FT		/tag= f	
FT		/product= "CstII sialyltransferase"	
FT		/note= "Open reading frame 7a"	
FT	CDS	6924..7964	
FT		/tag= g	
FT		/product= "sialic acid synthase"	
FT		/note= "Open reading frame 8a"	
FT	CDS	8020..9079	
FT		/tag= h	
FT		/product= "Sialic acid biosynthetic enzyme"	
FT		/note= "Open reading frame 9a"	
FT	CDS	9076..9741	
FT		/tag= i	
FT		/product= "CMP-sialic acid synthetase"	

FT	CDS	/note= "Open reading frame 10a"
FT	complement (10554..11366)	
FT	/*tag= j	
FT	/product= "Lipo oligosaccharide biosynthetic enzyme"	
FT	/note= "Open reading frame 12a"	
XX		
XX	WO2000046379-A1.	
XX		
XX	10-AUG-2000.	
XX		
XX	01-FEB-2000; 2000WO-CA000086.	
XX		
PR	01-FEB-1999; 99US-0118213P.	
XX		
PR	31-JAN-2000; 2000US-00495406.	
XX		
PA	(CANA ) NAT RES COUNCIL CANADA.	
XX		
FI	Gilbert M, Wakarchuk WW;	
XX		
DR	WPI; 2000-524418/47.	
DR	P-PSDB; AAY97200, AAY97201, AAY97202, AAY97203, AAY97204, AAY97205,	
DR	AAY97206.	
XX		
FT	Novel glycosyltransferase polypeptides and polynucleotides useful for	
PT	biosynthesis of ganglioside and ganglioside mimics, as diagnostic	
PT	reagents and as immunogen for producing antibodies.	
XX		
PS	Claim 1; Page 86-90; 120pp; English.	
XX		
CC	A reaction mixture for the synthesis of a sialylated oligosaccharide is	
CC	useful for synthesising sialylated oligosaccharide such as ganglioside,	
CC	lyso ganglioside or their mimics. Glycosyltransferases are useful for	
CC	chemo-enzymatic synthesis of oligosaccharides, including gangliosides and	
CC	other oligosaccharides that have biological activity. The enzymes and	
CC	nucleic acids that encode them are useful for studies of the pathogenesis	
CC	mechanisms of organisms that synthesize ganglioside mimics, such as C.	
CC	jejuni and the nucleic acids are used as probes to study expression of	
CC	genes involved in ganglioside mimetic synthesis. Antibodies raised	
CC	against the glycosyltransferases are also useful for analyzing the	
CC	expression patterns of these genes involved in pathogenesis. The nucleic	
CC	acids are also useful for designing antisense oligonucleotides for	
CC	inhibiting expression of the Campylobacter enzymes that are involved in	
CC	the biosynthesis of ganglioside mimics that can mask the pathogens from	
CC	the host's immune system. The oligosaccharides are useful as diagnosing	
CC	reagents or as therapeutics and as immunogens for producing antibodies.	
CC	Bacterial glycosyltransferase can be used to catalyse the formation of	
CC	oligosaccharides that are identical to the corresponding mammalian	
CC	structures and are easier and less expensive to produce in large	
CC	quantity, compared to the mammalian glycosyltransferase. The bacterial	
CC	origin of the enzymes facilitates expression of large quantities of the	
CC	enzymes using relatively inexpensive prokaryotic expression systems.	
CC	(Updated on 15-SEP-2003 to standardise OS field)	
XX		
SQ	Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;	
	Query Match 95.6%; Score 834.6; DB 3; Length 11474;	
	Best Local Similarity 97.3%; Pred. No. 5.6e-127;	
	Matches 849; Conservative 0; Mismatches 24; Indels 0; Gaps 0;	
QY	1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTC AAGG 60	
Db	6048 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGGATTATTC AAGA 6107	
QY	61 CTACCAAAATGATTTTGATGTATTAGTAGTAAATCAATTTTATTTTGAAGATAAATAC TAT 120	
Db	6108 CTACCAAAATGATTTTGATGTATTAGTAGTAAATCAATTTTATTTTGAAGATAAATAC TAT 6167	
QY	121 CTTGGTAAAAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC TAC 180	
Db	6168 CTTGGTAAAAAATGCAAGGAGTATTTTACATCTTATTCCTTTTTTTTGGACAAATAC TAC 6227	
QY	181 ACTTTTAAACATTTAATCCAAAATCAAGAAATAGAGCCGAACCTAAATTTATGTGTTCTTAAT 240	



CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyze the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 15-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;

Query Match		95.4%;	Score 833;	DB 3;	Length 876;
Best Local Similarity		97.1%;	Pred. No. 1e-126;		
Matches 848;		Conservative	0;	Mismatches 25;	Indels 0; Gaps 0;
QY	1	ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60		
DB	1	ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA	60		
QY	61	CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT	120		
DB	61	CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT	120		
QY	121	CTTGTTAAAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTCTTGAGCAATACTAC	180		
DB	121	CTTGTTAAAAATGCAAGGAGTATTTTACATCTCTATTTCTTTTGTGACAACTACTAC	180		
QY	181	ACTTTAAACATTTAATCCAAAATCAAGAATATGACCGAATTAATGTTGTTCTAAT	240		
DB	181	ACTTTAAACATTTAATCCAAAATCAAGAATATGACCGAATTAATGTTGTTCTAAT	240		
QY	241	TACACCAAGCTCATCAGAAAATGAAATTTTGTAAAACTTTTACGATTAATTTCTCT	300		
DB	241	TACACCAAGCTCATCAGAAAATGAAATTTTGTAAAACTTTTACGATTAATTTCTCT	300		
QY	301	GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	360		
DB	301	GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTAA	360		
QY	361	TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTCGAGTAGCC	420		
DB	361	TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTTATATGTCGAGTAGCC	420		
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAAATGGGTCA	480		
DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAAATGGGTCA	480		
QY	481	TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTCGATTTTAAAT	540		
DB	481	TCCTATGCTTTTGATACCAACAGAAATCTTTTAAACTAGCCCTCGATTTTAAAT	540		
QY	541	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTGAATTTCTA	600		
DB	541	GATTAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTGAATTTCTA	600		
QY	601	GAATAAACTTACAAAAATAAATCTATTTGCTTATGCTCTTATAGCTTTTACGAAATTTT	660		
DB	601	GAATAAACTTACAAAAATAAATCTATTTGCTTATGCTCTTATAGCTTTTACGAAATTTT	660		
QY	661	ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATAAGAAAAATAAATCACTACT	720		
DB	661	ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATAAGAAAAATAAATCACTACT	720		

QY	721	AAAGATATATCTCATCTCTAGCTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT	780
DB	721	AAAGATATATCTCATCTCTAGCTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTTT	780
QY	781	AAAAAATAAAAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA	840
DB	781	AAAAAATAAAAAATTTAAAGAAATATTTATTACAAAGTTGATAAAGATCTATTAAAGATTA	840
QY	841	CCTAGTGATATAAAGCATTTATTTTCAAAGGAAAA	873
DB	841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
RESULT 12			
ABT13666			
ID	ABT13666 standard; DNA; 876 BP.		
XX			
AC	ABT13666;		
XX			
DT	07-FEB-2003 (first entry)		
DE	C. jejuni bifunctional sialtransferase cstII coding sequence #1.		
XX			
KW	Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;		
KW	GalNAc transferase; N-Acetylglucosamine transferase;		
KW	galactosyltransferase; sialyltransferase; sialic acid synthase;		
KW	cytidine 5'-monophosphate sialic acid synthetase;		
KW	CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;		
XX	ganglioside mimetics; inflammation; tumour metastasis.		
OS	Campylobacter jejuni.		
XX			
PN	WO200274942-A2.		
XX			
PD	26-SEP-2002.		
XX			
PF	22-FEB-2002; 2002WO-CA000229.		
XX			
PR	21-MAR-2001; 2001US-00816028.		
XX			
PA	(CANA ) NAT RES COUNCIL CANADA.		
XX			
PI	Gilbert M, Wakarchuk WW;		
XX			
DR	WPI; 2003-040554/03.		
DR	P-PSDB; ABU18479.		
XX			
PT	New glycosyltransferases from Campylobacter, useful for synthesizing		
PT	gangliosides and ganglioside mimetics, and in studying the pathogenesis		
PT	mechanisms of organisms that synthesize ganglioside mimetics.		
XX			
PS	Claim 8; Page 96; 107pp; English.		
XX			
CC	The invention comprises the amino acid and coding sequences of		
CC	Campylobacter jejuni proteins. The C. jejuni proteins of the invention		
CC	may be either an: acyltransferase; glycosyltransferase; GalNAc (N-		
CC	Acetylglucosamine) transferase; galactosyltransferase;		
CC	sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)		
CC	sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein		
CC	sequences of the invention are useful for ganglioside synthesis, studying		
CC	ganglioside mimetics, and for designing oligonucleotides to inhibit		
CC	expression of Campylobacter enzymes involved in the biosynthesis of		
CC	ganglioside mimetics that can mask the pathogen's from the host's immune		
CC	system. The C. jejuni oligosaccharides of the invention may be used as		
CC	diagnostic reagents (e.g. to locate areas of inflammation or tumour		
CC	metastasis). The present DNA sequence represents a Campylobacter jejuni		
CC	gene of the invention		
XX			
SQ	Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;		
Query Match			
Best Local Similarity 97.1%; Score 833; DB 8; Length 876;			
Matches 848; Conservative 0; Mismatches 25; Indels 0; Gaps 0;			



```
Qy 1 ATGAAAAAGCTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60
Db 1 ATGAAAAAGCTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60
Qy 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGAGATAAATACATAT 120
Dy 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGAGATAAATACATAT 120
Qy 121 CTTGGTAAAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATACATAC 180
Dy 121 CTTGGTAAAAATGCAAGGAGTATTTTACAATCCTATCTCTTTTGTGAACATACTAC 180
Qy 181 ACTTTAAACATTTTAATCCAAAATCAAGATATGAGACCGAAGCTAAATATGTTCTTAAT 240
Dy 181 ACTTTAAACATTTTAATCCAAAATCAAGATATGAGACCGAAGCTAAATATGTTCTTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTGAAAACTTTTACGATATTATTTTCCCT 300
Dy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTCTGAAAACTTTTACGATATTATTTTCCCT 300
Qy 301 GATGCTCAATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Dy 301 GATGCTCAATTTGGGATATGATTTTAAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTTGCGAGTAGCC 420
Dy 361 TTTTACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
Dy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTAAAACTAGCCCTGATTTTAAAAAT 540
Dy 481 TCTTATGCTTTTGATACCAACAAAGAAATTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy 541 GATGCTTACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
Dy 541 GATAATTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTCTTATGCTTTTACGAAATTTT 660
Dy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTTCTTATGCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATCAAGAAAAATAACTACACT 720
Dy 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATCAAGAAAAATAACTACACT 720
Qy 721 AAAGATATCTCATACCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780
Dy 721 AAAGATATCTCATACCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780
Qy 781 AAAAAATAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAGATTA 840
Dy 781 AAAAAATAAAATTAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAGATTA 840
Qy 841 CCTAGTGATATAAGCATTTTTCAAAGGAAAA 873
Dy 841 CCTAGTGATATAAGCATTTTTCAAAGGAAAA 873
```

RESULT 13

AAAS3725

ID AAAS3725 standard; DNA; 876 BP.

XX AAAS3725;

XX 15-SEP-2003 (revised)

DT 22-DEC-2000 (first entry)

DE Campylobacter jejuni O:41 serotype alpha-2,3-sialyltransferase.

XX

Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
Beta-1,3-Galactosyltransferase; alpha-2,3-sialyltransferase;  
sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
immunity; immunogen; ganglioside; ds.

Campylobacter jejuni; O:41 serotype.

Key Location/Qualifiers  
CDS 1..876  
/\*tag= a  
/product= "alpha-2,3-sialyltransferase"

WO200046379-A1.

10-AUG-2000.

01-FEB-2000; 2000WO-CA000086.

01-FEB-1999; 99US-0118213P.

31-JAN-2000; 2000US-00495406.

(CANA ) NAT RES COUNCIL CANADA.

Gilbert M, Wakarchuk MW;

WPI; 2000-524418/47.

P-PSDB; AAY97211.

Novel glycosyltransferase polypeptides and polynucleotides useful for  
biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
reagents and as immunogen for producing antibodies.

Claim 6; Page 94; 120pp; English.

A reaction mixture for the synthesis of a sialylated oligosaccharide is  
useful for synthesising sialylated oligosaccharide such as ganglioside,  
lycoganglioside or their mimics. Glycosyltransferases are useful for  
chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
other oligosaccharides that have biological activity. The enzymes and  
nucleic acids that encode them are useful for studies of the pathogenesis  
mechanisms of organisms that synthesize ganglioside mimics, such as C.  
jejuni and the nucleic acids are used as probes to study expression of  
genes involved in ganglioside mimetic synthesis. Antibodies raised  
against the glycosyltransferases are also useful for analyzing the  
expression patterns of these genes involved in pathogenesis. The nucleic  
acids are also useful for designing antisense oligonucleotides for  
inhibiting expression of the Campylobacter enzymes that are involved in  
the biosynthesis of ganglioside mimics that can mask the pathogens from  
the host's immune system. The oligosaccharides are useful as diagnosing  
reagents or as therapeutics and as immunogens for producing antibodies.  
Bacterial glycosyltransferase can be used to catalyze the formation of  
oligosaccharides that are identical to the corresponding mammalian  
structures and are easier and less expensive to produce in large  
quantity, compared to the mammalian glycosyltransferase. The bacterial  
origin of the enzymes facilitates expression of large quantities of the  
enzymes using relatively inexpensive prokaryotic expression systems.  
(Updated on 15-SEP-2003 to standardise OS field)

Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Query Match 93.6%; Score 817; DB 3; Length 876;  
Best Local Similarity 96.0%; Pred. No. 4.2e-124;  
Matches 838; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60

Dy 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60

Qy 61 CTACCAATGATTTTGGTGTATTAGATGTAATCAATTTTATTGAAAGATAAATCTAT 120

Dy 61 CTACCAATGATTTTGGTGTATTAGATGTAATCAATTTTATTGAAAGATAAATCTAT 120



QY 121 CTTGGTAAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180  
 Db 121 CTTGGTAAATAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTCTTTGAGCAATAC 180  
 QY 181 ACTTTAAACATTTTAATCAAAATCAAGAATATGAGCCGAACTAAATATGTGTTCTAAT 240  
 Db 181 ACTTTAAACATTTTAATCAAAATCAAGAATATGAGCCGAACTAAATATGTGTTCTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTAAACCTTTTACGATTATTTTCT 300  
 Db 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTAAACCTTTTACGATTATTTTCT 300  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATCTTATTTTAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATCTTATTTTAA 360  
 QY 361 TTTACGAAATTTATTTCAATCAAGAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420  
 Db 361 TTTACGAAATTTATTTCAATCAAGAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
 Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
 QY 541 GATCCTCAGCTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAAATTTCTA 600  
 Db 541 GATCCTCAGCTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAAATTTCTA 600  
 QY 601 GAAAAAATTTTACAAATAAATCTATATGCTTTTCTTATGCTTCTTAAAGCTTTTAAATTTT 660  
 Db 601 GAAAAAATTTTACAAATAAATCTATATGCTTTTCTTATGCTTCTTAAAGCTTTTAAATTTT 660  
 QY 661 ATAGAACTAGCCGCAAAATTTAAATCAAAATTTTATCATACAAAGAAAATTAACACT 720  
 Db 661 ATAGAACTAGCCGCAAAATTTAAATCAAAATTTTATCATACAAAGAAAATTAACACT 720  
 QY 721 AAAGATATCTATACATCTTCTAGTGGCTTATGAGAAATTTTCAAAATTTTAAATTTT 780  
 Db 721 AAAGATATCTATACATCTTCTAGTGGCTTATGAGAAATTTTCAAAATTTTAAATTTT 780  
 QY 781 AAAAAATATAAATAAAGAAATCTTATTAACAAGTTGATATAAGATCTATTAAAGATTA 840  
 Db 781 AAAAAATATAAATAAAGAAATCTTATTAACAAGTTGATATAAGATCTATTAAAGATTA 840  
 QY 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873  
 Db 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873

RESULT 14

ABT13668

ID ABT13668 standard; DNA; 876 BP.

XX AC ABT13668;

XX DX 07-FEB-2003 (first entry)

XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #3.

XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;

XX KW GalNAc transferase; N-Acetylgalactosamine transferase;

XX KW Galactosyltransferase; sialyltransferase; sialic acid synthase;

XX KW cytidine 5'-monophosphate sialic acid synthetase;

XX KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;

XX KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.

XX XX WO200274942-A2.

XX FN

XX 26-SEP-2002.  
 XX 22-FEB-2002; 2002WO-CA000229.  
 XX 21-MAR-2001; 2001US-00816028.  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 XX Gilbert M, Wakarchuk WW;  
 XX WPI: 2003-040554/03.  
 XX P-PSDB; ABJ18481.  
 XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 XX mechanisms of organisms that synthesize ganglioside mimetics.  
 XX Claim 8; Page 97; 107pp; English.  
 XX The invention comprises the amino acid and coding sequences of  
 XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 XX may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 XX Acetylgalactosamine) transferase; galactosyltransferase;  
 XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 XX sequences of the invention are useful for ganglioside synthesis, studying  
 XX ganglioside mimetics, and for designing oligonucleotides to inhibit  
 XX expression of Campylobacter enzymes involved in the biosynthesis of  
 XX ganglioside mimetics that can mask the pathogen's from the host's immune  
 XX system. The C. jejuni oligosaccharides of the invention may be used as  
 XX diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 XX metastasis). The present DNA sequence represents a Campylobacter jejuni  
 XX gene of the invention  
 XX Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;  
 QY Query Match 93.6%; Score 817; DB 8; Length 876;  
 Db Best Local Similarity 96.0%; Pred. No. 4.2e-124;  
 QY Matches 838; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
 Db 1 ATGAAAAAGTATTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
 QY 1 ATGAAAAAGTATTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
 Db 61 CTACCAAAATGATTTGATGTTTATAGATGTAATCAATTTTATTTGAGTAATACTAT 120  
 QY 61 CTACCAAAATGATTTGATGTTTATAGATGTAATCAATTTTATTTGAGTAATACTAT 120  
 Db 121 CTTGGTAAAAAATGCAAAAACAGTGTGTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
 QY 121 CTTGGTAAAAAATGCAAAAACAGTGTGTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
 Db 181 ACTTTAAAAATTTAATCCAAATCAAGAATATGAGACCGAACTAAATCATGTGTTCTAAT 240  
 QY 181 ACTTTAAAAATTTAATCCAAATCAAGAATATGAGACCGAACTAAATCATGTGTTCTAAT 240  
 Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTATTTTCT 300  
 QY 241 TTTAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTATTTTCT 300  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATCTTATTTTAA 360  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACCACTTAAAGAAATTAATCTTATTTTAA 360  
 Db 361 TTTACGAAATTTTATTTCAATCAAGAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420  
 QY 361 TTTACGAAATTTTATTTCAATCAAGAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Db



Search completed: July 19, 2006, 15:26:24  
Job time : 656.252 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:59:51 ; Search time 4154.24 Seconds  
(without alignments)  
11751.261 Million cell updates/sec

Title: US-10-734-719-13

Perfect score: 873  
Sequence: 1 atgaaaaagttatttcgc.....agcattatttcaaggaaaa 873

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1.\*  
2: gb\_est3.\*  
3: gb\_est4.\*  
4: gb\_est5.\*  
5: gb\_est6.\*  
6: gb\_est7.\*  
7: gb\_est8.\*  
8: gb\_est9.\*  
9: gb\_est10.\*  
10: gb\_est11.\*  
11: gb\_est12.\*  
12: gb\_est13.\*  
13: gb\_est14.\*  
14: gb\_est15.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	93.6	10.7	1359	9	DN685273
C 2	93.4	10.7	1101	14	AL069706 Drosophil
C 3	92	10.5	1365	10	DT959930 CFM165-H0
C 4	91.6	10.5	1227	14	AG430010 Mus muscu
C 5	90.6	10.4	1101	14	CNS0039G
C 6	90.6	10.4	1179	1	AJ927794
C 7	90	10.3	1007	9	DN566957
C 8	87.4	10.0	1354	12	CG744717
C 9	87.2	10.0	976	13	CW949642
C 10	86.6	9.9	1210	12	CG749728
C 11	86.4	9.9	1542	14	AG386981
C 12	85.8	9.8	1096	1	AJ927821
C 13	85	9.7	1364	12	CG757970
C 14	84.6	9.7	1242	12	CL068807
C 15	84.4	9.7	1101	14	CNS00B01
C 16	84.4	9.7	1324	14	AG376784
C 17	84.4	9.7	1489	10	DT957227
C 18	83.8	9.6	1297	10	DV778934
C 19	83.6	9.6	734	14	CNS010MP

C 20	83.6	9.6	1123	5	CD325190
C 21	83.6	9.6	1227	10	DT989558
C 22	83.6	9.6	1355	14	AG346348
C 23	83.4	9.6	906	14	AG555384
C 24	83.2	9.5	1277	12	CC253231
C 25	83.2	9.5	1544	10	DV780516
C 26	83	9.5	930	7	BB898645
C 27	83	9.5	1092	1	AJ927993
C 28	83	9.5	1258	13	CL508886
C 29	83	9.5	1268	14	AG347098
C 30	82.8	9.5	1313	8	CK997149
C 31	82.8	9.5	1348	12	CG749499
C 32	82.6	9.5	1248	9	DN808090
C 33	82.4	9.4	947	12	CL112501
C 34	82.2	9.4	963	10	DT797234
C 35	82.2	9.4	1101	14	CNS00EVL
C 36	82.2	9.4	1118	10	DM600040
C 37	82.2	9.4	1186	13	CL649642
C 38	82	9.4	1638	14	AG347262
C 39	81.8	9.4	964	13	CW957121
C 40	81.6	9.3	759	14	CNS06QXV
C 41	81.6	9.3	949	5	CK424488
C 42	81.6	9.3	1592	12	CG750135
C 43	81.6	9.3	1626	5	CF238805
C 44	81.2	9.3	885	9	DN571958
C 45	81.2	9.3	1028	13	CW954747

## ALIGNMENTS

### RESULT 1

DN685273/c

LOCUS

DEFINITION

CGX40-G12.yid-s SHGC-CGX Gasterosteus aculeatus cDNA clone

CGX40-G12 5', mRNA sequence.

DN685273

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

DN685273 1359 bp mRNA linear EST 30-MAR-2005  
CGX40-G12.yid-s SHGC-CGX Gasterosteus aculeatus cDNA clone  
CGX40-G12 5', mRNA sequence.  
DN685273 GI:62033458  
EST.  
Gasterosteus aculeatus (three spined stickleback)  
Gasterosteus aculeatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;  
Gasterosteidae; Gasterosteus.  
1 (bases 1 to 1359)  
Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,  
Schmutz,J. and Myers,R.M.  
Expressed sequence tags from Gasterosteus aculeatus  
Unpublished (2003)  
Contact: Grimwood, Jane  
Stanford Human Genome Center  
Stanford University School of Medicine  
975 S California Ave, Palo Alto, CA 94304, USA  
Tel: 650 320 5917  
Fax: 650 320 5801  
Email: janes@hgc.stanford.edu  
Plate: 40  
High quality sequence start: 18  
High quality sequence stop: 102.  
Location/Qualifiers  
1..1359  
/organism="Gasterosteus aculeatus"  
/mol\_type="mRNA"  
/strain="Comner Creek sticklebacks, WA USA"  
/db\_xref="taxon:69293"  
/clone="CGX40-G12"  
/sex="mixed male and female"  
/tissue\_type="eyes"  
/dev\_stage="adult"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="SHGC-CGX"  
/note="Vector: Express 1; Total and poly A+ RNA was



Qy	453	GGGAATTCATTTTTTATCAAAATGGTCATCTTATGCTTTTCATACCAAAACAAGAAAATCT	512
Db	687	ATAWAAAAAATAAAATWAAWAATWATATAAATTTAAALAAWAATAAAAWAAWAAT	746
Qy	513	TTTAAAACTAGCCCTGATTTTAAAAATGATCGCTCACACTATATCGGACACTAGTAAAAA	572
Db	747	WAWWATAATATWATATATATATWTTWAWWATAWAAWMMTA-TAWATAWAATAWAAWAWAWA	805
Qy	573	TACAGATATAAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACTATATTCGCT	632
Db	806	TAAATATATATWAAAWAAWATAWATWATATATAWAAATAWAAAAAAATWTAATATWA	865
Qy	633	ATGTCCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCAAAATTTAAATTCAAAATTT	692
Db	866	ATWATAWAAAAATAWAWTTTWTWTTTWWAWATATAAAWAWATAWAAWAAAAAATAAAA	925
Qy	693	TATCATACAGAAAAAATAACTACACATAAGATATATCTCATCTTCTAGTGGAGCTTA	752
Db	926	AAAAATAAWAWWTTWATTTTATTAAAWTTWATWATTTWATTTWAAWTTWATATWTTTWA	985
Qy	753	TGGAAAAATTTTCAAAAAATTAATTTTAAAAAATAAAAAATTAAGAAAAATGTTTATTA	812
Db	986	TWTATATATWWTAWATATATTTATTAAWWTATATTTTAWAAWTAATATATWATAWW	1045
Qy	813	CAAGTTGCATAAAGATCTATTAGATTACCTAGTGATATAAGCATTTATTTCAA	867
Db	1046	WTAWATATAWAWWAATTAATTTATATATWTAAWATAWMAAAWAAAAWAAWATAWMA	1100

RESULT 3					
DT959930/c					
LOCUS	DT959930	1365 bp	mrna	linear	EST 21-SEP-2005
DEFINITION	CFW165-H01.x1d-t SHGC-CFW Gasterosteus aculeatus cdna clone				
	CFW165-H01 3', mRNA sequence.				

## ORIGIN

Query Match	10.5%	Score 92	DB 10	Length 1365
Best local Similarity	45.7%	Pred. No. 1.6e-05		
Matches	384	Conservative 0	Mismatches 446	Indels 11
Gaps	2			
QY	38	TAAAAGAAATTGATTATTCAAGCTACCAATGATTTTGATGTATTAGATGTAATCAAT	97	
DB	1292	TATAAAAAATTAAATAATAATTTTATATTAATAAAAAATATTTAAAAAATAATTTTAA	1233	
QY	98	TTTATTTTGAAGATAAATACTATCTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTA	157	
DB	1232	AATNTAATTTNTTAAATTAATAATAATTAATAATTTATATATTTTATAATAATAA	1173	
QY	158	ATTTCCTCTTTGAGCAATACTACACCTTAAACATTTAATCCAAAATCAAGATAATGAGA	217	
DB	1172	TATAAATTTTAAAAAATAATTTTAAATAATAAAAAATTTATATATAATATATATT	1113	
QY	218	CGAAGCTAAATTTGTGTTCTTAATTACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAA	277	
DB	1112	TAAATATATTTAATAATTAATNATAAAAAAATAATTTATTTAAAAATAATATATATTTTA	1053	
QY	278	AAACTTTTTACGATTAATTTTCTGATGCTCATTTGGGATATAGATTTTTTTTAAACAACCTTA	337	
DB	1052	ATATTAATTTAAAAATAAATAATATATAAAAAATAAAAAATAATAAAATAATTTATTTT	993	
QY	338	AGAATTTTAATGCTTAATTTTAAATTTTCAGAAATTTATTTTCAATCAAGAAATTTACCTCAG	397	
DB	992	AAAATAATAAAAAATATTTTATNTTATTTTATATAAAAAATTTATTTAAAAATTTAATAAT	933	
QY	398	GGGTCTATATGTGCGAGTAGCATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAA	457	
DB	932	AAAAAAATATAAATAATAATTTTATAATAATAAATTTAAATAAATTTTATAATAATAA	873	
QY	458	TTGATTTTTATCAAAATGGTCACTCTATGCTTTTGATACCAAAACAAAGAAATCTTTTAA	517	
DB	872	TATATATTAATAATTTTATTTTAAATATAAATTTNATTTATTAATAAAAAAATATTTT	813	
QY	518	AACTAGCCCCGTATTTTAAAAATGATF-----CGCTCACACTATATCGACATAGTAAA	570	
DB	812	ATATATTTNTAAATTTTAAAAAATAATAATAATAATTAATTAATTAATAATAATAA	753	
QY	571	AATACAGATATAAAGCTTTGAAATTTCTAGAA-----AAAACTTCAAAATAAAACTATA	626	
DB	752	AATAATAATAAAAAAATATAAAATTTTTTAAATAATAAATTTAAATAATTTATAAAATA	693	
QY	627	TWGTCTTAGTCCCTAATAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTTAAATTC	686	
DB	692	TAAAAAATAATTTATTTAATAATAATAATAATAAAAAATAAAAAATTTTAAAAAATTTTAA	633	
QY	687	AAATTTTATCATACAGAAAAAATAAATACTACATAAGATATACTCATACCTTCTTAGTGA	746	
DB	632	ATATTTATTTTATATAAAAAATAATAATTTATTAATTTAAAAAATAATAATAATAATNAT	573	
QY	747	GGCTTATGAAAAATTTTCAAAAAAATTTTAAATTTTTTAAAAAATAAAAAATTTTAAAAAATGT	806	









TITLE An integrated physical and genetic map of the nematode *Pristionchus*  
JOURNAL Pacificus  
PUBMED Mol. Genet. Genomics 269 (5), 715-722 (2003)  
COMMENT 12884007  
Contact: Sommer RJ  
Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
Class: BAC ends.

FEATURES Location/Qualifiers  
source 1..1354  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="ppa EcoRI BAC Library"  
/notes="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

ORIGIN  
Query Match 10.0%; Score 87.4; DB 12; Length 1354;  
Best Local Similarity 46.8%; Pred. No. 9.5e-05;  
Matches 391; Conservative 0; Mismatches 436; Indels 8; Gaps 4;

QY 4 AAAAAAGTTATTATTCTCGGAATGACCAAGTTTAAAGAAATTCATTTATTCACAGGCTA 63  
DB 1319 AAAAAATTTTAAAAAATAATATAATNTAAATAATAATAATAATAATAATAATAAT 1260  
QY 64 CCAATGATTTTGATGATTTAGATGATCAATTTTATTTTGAAGATAAATCTATCTT 123  
DB 1259 AAAAAAATAATATATATTTATATATATATATATATATTTTATATATATATATAT 1200  
QY 124 GGTAAAAATGCAAAACAGTGTTTTACCCCTAAATTTCTCTTTGAGCAATCTACACT 183  
DB 1199 AATAAATAATATATTTAAAAATATATATTTTAAATATATTTTAAATATAATAAAT 1140  
QY 184 TTAACAATTTTAAATCAAGATATGACCGAATCAATTTATGTTCTAATATAC 243  
DB 1139 AATATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1084  
QY 244 AACCAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTTCTCTGAT 303  
DB 1083 AATAAATTTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATTTTAT 1024  
QY 304 GCTCAATTTGGGATGATTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTAAATTT 363  
DB 1023 ATAAATATATATATATAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 965  
QY 364 CACGAAATTTTATTCATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCATA 423  
DB 964 TATTATAAATTTTATAAATAATATATATAAATAAATAAATAAATAAATAAATAAATAA 905  
QY 424 GCCTTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTTTATCAAAATGGTCACT 483  
DB 904 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 845  
QY 484 TAAGCTTTGATACCAACAGAAATCTTTTAAACTAGCCCTCGATTTTAAATAATGAT 543  
DB 844 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 785  
QY 544 CGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTTAGAA 603  
DB 784 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 725  
QY 604 AAAAATTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 663  
DB 724 AATATTTATTCGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 665  
QY 664 GAACTAGCGCCAAATTTTAAAT-TCAAATTTTATCATACAGAAAAAATAAATAAATAA 722

DB 664 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 605  
QY 723 AGATATATCTACCTCTAGTGGCTTATGAAAAATTT--TCAAAAAATATTAATTTT 780  
DB 604 ATATTTATTTTATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 545  
QY 781 AAAAAATAAATAAATAAAGAAATGTTTATTAACAAGTTGATAAAAAGATCTATTA 835  
DB 544 ANTAAATTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 490  
RESULT 9  
CW949642 976 bp DNA linear GSS 20-DEC-2004  
LOCUS TCB34.3\_D04.SP6.Tribolium BAC library Tribolium castaneum genomic,  
DEFINITION genomic\_survey sequence.  
ACCESSION CW949642  
VERSION CW949642.1 GI:56732262  
KEYWORDS GSS.  
SOURCE Tribolium castaneum (red flour beetle)  
ORGANISM Tribolium castaneum  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;  
Tenebrionidae; Tribolium.  
REFERENCE 1 (bases 1 to 976)  
AUTHORS Savard, J. and Tautz, D.  
TITLE Tribolium castaneum BAC-ends sequencing project  
JOURNAL Unpublished (2003)  
COMMENT Contact: Savard, J.  
Abteilung für Evolutionsgenetik, AG Tautz  
Institut für Genetik, Universität zu Köln  
Weyertal 121, 50931 Köln, Germany  
Tel: 49 221 470 6911  
Fax: 49 221 470 5975  
Email: savard@uni-koeln.de  
Class: BAC ends.  
FEATURES Location/Qualifiers  
source 1..976  
/organism="Tribolium castaneum"  
/mol\_type="genomic DNA"  
/strain="GA-2"  
/db\_xref="taxon:7070"  
/clone\_lib="Tribolium BAC library"  
/notes="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;  
Library constructed by Exelixis Inc."  
ORIGIN  
Query Match 10.0%; Score 87.2; DB 13; Length 976;  
Best Local Similarity 44.0%; Pred. No. 0.00011;  
Matches 349; Conservative 0; Mismatches 441; Indels 4; Gaps 1;  
QY 84 TAGATGTAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAAATGCAAAACAGT 143  
DB 5 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 64  
QY 144 GTTTTACACCCCTAAATTTCTTCTTGGACCAATCTACACTTTAAACATTTAATCCAAA 203  
DB 65 ATAAGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 124  
QY 204 TCAAGAATATGAGACCGAACTAATTTATGTGTTCTTAATATACCAACCAAGCTCATAGAAA 263  
DB 125 TAAAAAAGAGTTNN 184  
QY 264 TGAATAATTTTGTAAAACTTTTTCACGATTTATTTTCTGATGCTCATTTGGGATATGATTT 323  
DB 185 TAAATAAGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 244  
QY 324 TTTTAAACACTTTAAAGAAATTTAATGCTTATTTTAAATTTTAAATTTTAAATTTTAAATCA 383  
DB 245 ATAAAAATTAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 304  
QY 384 AAGAATTACCTCAGGGGTCTATATGTGTGCGAGTAGCCATAGCCCTTAGGATACAAAGAAAT 443



```

AG386981.1 GI:47998186
VERSION Mus musculus molossinus (Japanese wild mouse)
KEYWORDS Mus musculus molossinus
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus

REFERENCE
AUTHORS Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
        Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriaki,K. and
        Shiroishi,T.
TITLE Contribution of Asian mouse subspecies Mus musculus molossinus to
        genomic constitution of strain C57BL/6J, as defined by BAC-end
        sequence-SNP analysis
JOURNAL Genome Res. 14 (12), 2439-2447 (2004)
PUBMED 15574823
REFERENCE
AUTHORS Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-NOV-2003) Mashira Hattori, The Institute of Physical
        and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
        1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
        (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
        Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSWg01. For BAC
        library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
        Tsukuba Institute, Bio Resource Center,
        The Institute of Physical and Chemical Research (RIKEN) 3-1-1
        Koyadai, Tsukuba, 305-0074 Japan
        phone: 81-298-36-9189, fax: 81-298-36-9199
        e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACE3.6
Vector : ECoRI
R.Site 1 : ECoRI.
R.Site 2 : ECoRI.
FEATURES
    source
        1..1542
            Location/Qualifiers
                1..1542
                    /organism="Mus musculus molossinus"
                    /mol_type="genomic DNA"
                    /sub_species="molossinus"
                    /db_xref="taxon:57486"
                    /clone="MSWg01-201G10.TJ"
                    /sex="male"
                    /tissue type="mixture of kidney and spleen"
                    /clone_lib="MSWg01 Mouse Male BAC Library"

ORIGIN
Query Match 9.9%; Score 86.4; DB 14; Length 1542;
Best Local Similarity 46.7%; Pred. No. 0.00014;
Matches 409; Conservative 0; Mismatches 452; Indels 14; Gaps 5;

QY 5 AAAAAAGTTATTTCCTGCTGGAATGACCAAGTTTAAAGAAATTTGATTTCACAGCTAC 64
Db 1441 ATNAAATAAATANTANTNATAATATATTAAGATAATAATAATATATATATATATAN 1382
QY 65 CAAATGATTTTCATGCTATTAGATGTAATCAATTTTATTTTGAAGATAATACTACTCTG 124
Db 1381 AAAAAATTAAATAATAATAAANNATAGTAATAATAATAATAATAATAATAATAATA 1322
QY 125 GTAAAAATGCAAAACAGTGTGTTTACACCCCTCAATTTCTCTTTGAGCAATCACTAC 184
Db 1321 ATTAATAT-ATATATATAATAATAAATAATATATATATTTANATAATAATAATAA 1263
QY 185 TAAACATTATCCAAATCAAGAAATATGAGACCGAAGCTAATTTGCTTCTTAATTACA 244
Db 1262 TAAATAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1206
QY 245 ACCAGCTCACTAGAAATGAAATTTGTAAAAACITTTTACGATTTATTTCTCGATG 304
Db 1205 AAATAATAATATATATATATATATATATTAATAATAATAATAATAATAATAATAA 1146

```

```

QY 305 CTCATTGGGATATGATTTTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTAAATTC 364
Db 1145 TTTAAATAAAATATATTTAAATAATAAAAAAAATAATANTAAATAAGAAAAATTTATATAAT 1086
QY 365 ACGAAATTTTATTCATCAAGAAATTTACCTCAGGGCTCTATATG-TGTGAGTAGGCATA 423
Db 1085 AAATANTATAAAAAATATAATTAATAATAAANTATTTAATAATAATAATAAANAATAAAA 1026
QY 424 GCCCTAGGATCAAAAGAAATTTATCTCTCGGAAATTTGATTTTTTATCAAAATGGTCACTCT 483
Db 1025 AAAATAAAATATAAAAAAAATAATTTTAAATAAANAATAATAATAATAATAATAATA 970
QY 484 TATGCTTTTGATACCAACAAGAAATCTTTTAAACATAGCCCTGATTTTAAAAATGAT 543
Db 969 TATAATATTTAATAAAATATAATATAATAAATAAANAATTTATAATTAATAATAATAAAT 910
QY 544 CGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGAACTTTAGAAATTTCTAGAA 603
Db 909 TATTAAAAAAATAAATAATTTATAAATAATATATATAATAATAATTTTAAATTTATAATA 850
QY 604 AAAAATTTACAAAAATAAACTATATTGCTTATGTCCTAATAGTCTTTTACGAAATTTTATA 663
Db 849 AAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 790
QY 664 GAATAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAA-----AAAAATAACTACA 718
Db 789 AAAATAAAAAAAATAAATAATAATAATAAATAAATAAATAAATAAATAAATAAATAATA 730
QY 719 CTAAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAAT 778
Db 729 ATAAATAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATATAT 670
QY 779 TTAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 838
Db 669 TTATAATAAATAATAATAATAATAATAATAATAATAATAAATAAATAAATAAATAAATAA 610
QY 839 TACCTAGTATATAAAGCATTATTTCACAAAGGAAA 873
Db 609 AATATTTAATATTATTAAATNTAATTATTAAAAA 575

RESULT 12
AJ927821 1096 bp mRNA linear EST 10-JUN-2005
LOCUS AJ927821 Theileria annulata merozoite Theileria annulata cDNA clone
DEFINITION tam027c04_glk, mRNA sequence.
ACCESSION AJ927821
VERSION AJ927821.1 GI:67498206
KEYWORDS EST.
SOURCE Theileria annulata
ORGANISM Theileria annulata
Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae; Theileria.
REFERENCE
AUTHORS 1 (bases 1 to 1096)
        Pain,A., Renauld,H., Berriman,M., Murphy,L., Yeats,C.A., Weir,W.,
        Kerhornou,A., Aslett,M., Bishop,R., Bouchier,C., Cochet,M.,
        Coulson,R.M.R., Cronin,A., de Villiers,E., Fraser,A., Fosker,N.,
        Gardner,M., Gobler,A., Griffiths-Jones,S., Harris,D.E., Katzer,F.,
        Larke,N., Lord,A., Maser,P., McKellar,S., Mooney,P., Morton,F.,
        Nene,V., O'Neill,S., Price,C., Quail,M.A., Rabinowitsch,E.,
        Rawlings,N.D., Rutter,S., Saunders,D., Seeger,K., Shah,T.,
        Squares,R., Squares,S., Tivey,A., Walker,A.R., Woodward,J.,
        Dobbelaere,D.A.E., Langsley,G., Rajandream,M.-A., McKeever,D.,
        Shiels,B., Tait,A., Barrell,B. and Hall,N.
TITLE The genome of the host-cell transforming parasite Theileria
        annulata and a comparison with T. parva
JOURNAL Unpublished (2005)
COMMENT Contact: Pain A
        The Pathogen Sequencing Unit
        The Wellcome Trust Sanger Institute
        Genome Campus, CB10 1SA, UNITED KINGDOM
        Merozoite cDNA library: Frank Katzer and Brian Shiels, Division of
        Veterinary Infection and Immunity, ICM, University of Glasgow, UK.

```

FEATURES	Location/Qualifiers	KEYWORDS
source	1..1096	GSS.
	/organism="Theileria annulata"	Pristionchus pacificus
	/mol_type="mRNA"	Pristionchus pacificus
	/isolate="Ankara (clone D7)"	Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
	/db_xref="taxon:5874"	Neodiplogasteridae; Pristionchus.
	/clone="cam027c04_qlk"	1 (bases 1 to 1364)
	/dev_stage="merozoite"	Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
	/lab_host="Bos taurus (cow)"	Buntjer,J., van der Meulen,M. and Sommer,R.J.
	/clone_lib="Theileria annulata merozoite"	An integrated physical and genetic map of the nematode Pristionchus
	/note="country: Turkey;Ankara"	pacificus
ORIGIN		Mol. Genet. Genomics 269 (5), 715-722 (2003)
	Query Match 9.8%; Score 85.8; DB 1; Length 1096;	12884007
	Best Local Similarity 36.7%; Pred. No. 0.00018;	Contact: Sommer RJ
	Matches 262; Conservative 0; Mismatches 446; Indels 6; Gaps 1;	Evolutionary Biology
QY	160 TTCTCTTTGAGCAATACACCTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACC 219	Max-Planck-Institute for Developmental Biology
DB	307 TTTTNNNNNGGANNNNNATTANNNNNAAGTATNNNNNNNAAGTAATTAATAAAAAAN 366	Spenannstr. 37-39, Tuebingen D-72076, Germany
QY	220 GAACCTAATTATGTCTTAATACAAACCAAGCTCATCTAGAAAAATGAAAAATTTGTA 279	Tel: 00497071601371
DB	367 ANANNAANGTTAATGTTTAATATANNANNGNNNTTATTTTAATAATANNAAATNTG 426	Fax: 00497071601498
QY	280 ACTTTTTCAGATTTATTTCTGATGCTCATTTGGGATATGATTTTAAAAACAATTTAA 339	Email: ralf.sommer@tuebingen.mpg.de
DB	427 NNTAATATANATNTAATATATAAATNNNTNTATTTATATTANAAAAAANNNTNNNN 486	Class: BAC ends.
QY	340 GAATTTAATGCTTATTTTAAATTTTCAGAAATTTATTTCAATCAAGAAATACCTCAGG 399	FEATURES
DB	487 NNANAAAAAGTAAATTAATAAAAAAANNNNNTGNTTAAANATTAATAATAATAATAA 546	Location/Qualifiers
QY	400 GTCTATATGTCGAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAAT 459	source
DB	547 NNNAAAAATTAATANNNAANAAAAAANNAAAAAANNAAAAAANNAAAAATTTTAAAT 606	1..1364
QY	460 GATTTTATCAAAATGGGTCATCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAA 519	/organism="Pristionchus pacificus"
DB	607 TAANTATATAATANNAAAAATTTAANTAANNAANNAANNAANNAANNAANNAANNA 666	/mol_type="genomic DNA"
QY	520 CTAGCCCTGATTTTAAAAATGATCGCTCACACTATATCGGACATAGTAAAAATACAGAT 579	/strain="Californica"
DB	667 TTANAAAA-----ANAAAAAANNAAAAAATTTTANAAAAAATNNNTNNNAAAN 720	/db_xref="taxon:54126"
QY	580 ATAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAAGAAATTTGCTTATGCTCT 639	/clone_lib="Ppa EcoRI BAC Library"
DB	721 NATANNNAANNNNNNANNAATAAANNAATAAATAAATAAATAAATAAATAAATAAATA 780	/note="The library was generated by a partial digest of
QY	640 AATAGTCTTTTATAGCAATTTTATAGCACTAGCGCCAAATTTTAAATTTCAAAATTTTAT 699	the genomic DNA with EcoRI and cloning into the BAC
DB	781 AANNATATTTANAANNTTTTAAANNNAANNAANNAANNAANNAANNAANNAANNAANNA 840	vector."
QY	700 CAAGAAAAATAACTACACTAAAGATATATCTATACCTCTAGTGAGGCTTATGGAATA 759	ORIGIN
DB	841 AAAAAAANTANNATNTNNNAATTTANNATNAAAAATNNAAAAAATAGTATTTGTANTAAA 900	Query Match 9.7%; Score 85; DB 12; Length 1364;
QY	760 TTTTCAAAAAATATTAATTTTAAAAAATAAATAAATAAAGAAATGCTTTTATACAGTTG 819	Best Local Similarity 44.9%; Pred. No. 0.00024;
DB	901 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960	Matches 391; Conservative 0; Mismatches 464; Indels 16; Gaps 2;
QY	820 ATAAAGATCTATTAAAGATTACCTAGTGTATATAAGCAATTTTCAAAAGGAAAA 873	QY 4 AAAAAAGTTATTATTGCTCGAAATGACCAAGTTTAAAGAAATGATTGATTCAAGGCTA 63
DB	961 ANNAANATATANAAAAAATAAATAATAAATAATAAATAAATAAATAAATAAATAAATAA 1014	DB 1324 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1265
RESULT 13		QY 64 CCAATGATTTTGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATCTATCTT 123
CG757970/c		DB 1264 TATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1205
LOCUS	1364 bp DNA linear GSS 24-OCT-2003	QY 124 GGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTCTTTCGAGCAATCTACACT 183
DEFINITION	P053-2-C06.zb Ppa EcoRI BAC Library Pristionchus pacificus genomic,	DB 1204 TTTTAAAAATTTATTTATTTTNNATATANAATAATAATAATAATAATAATAATA 1145
ACCESSION	genomic survey sequence.	QY 184 TTAATAACATTTAATCCAAATCAAGAAATATGACCGCACTAATTTATGTTGTTCTAATTAC 243
VERSION	CG757970	DB 1144 ATTAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1085
	CG757970.1 GI:37987066	QY 244 AACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTACGATTTATTTTCCTGAT 303





GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:02:20 ; Search time 136.177 Seconds  
(without alignments)  
11995.226 Million cell updates/sec

Title: US-10-734-719-13  
Perfect score: 873  
Sequence: 1 atgaaaaagttattatgc.....agcattatttcaaggaaaaa 873

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/ina/1\_COMB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/ina/5\_COMB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/ina/7\_COMB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/ina/H\_COMB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/ina/PP\_COMB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE\_COMB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/ina/RE\_COMB.seq.\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	873	100.0	873	3	US-09-816-028A-13
2	873	100.0	873	3	US-10-303-162-13
3	873	100.0	873	3	US-10-303-134-13
4	873	100.0	873	3	US-10-303-118-13
5	873	100.0	873	3	US-10-303-128-13
6	873	100.0	873	5	US-10-735-419-13
7	869.8	99.6	876	3	US-09-495-406-8
8	869.8	99.6	876	3	US-09-816-028A-8
9	869.8	99.6	876	3	US-10-303-162-8
10	869.8	99.6	876	3	US-10-303-134-8
11	869.8	99.6	876	3	US-10-303-118-8
12	869.8	99.6	876	3	US-10-303-128-8
13	869.8	99.6	876	5	US-10-735-419-8
14	866.6	99.3	873	3	US-09-816-028A-11
15	866.6	99.3	873	3	US-10-303-162-11
16	866.6	99.3	873	3	US-10-303-134-11
17	866.6	99.3	873	3	US-10-303-118-11
18	866.6	99.3	873	3	US-10-303-128-11
19	866.6	99.3	873	5	US-10-735-419-11
20	841	96.3	876	3	US-09-495-406-4
21	841	96.3	876	3	US-09-816-028A-4
22	841	96.3	876	3	US-10-303-162-4
23	841	96.3	876	3	US-10-303-134-4

24	841	96.3	876	3	US-10-303-118-4	Sequence 4, Appli
25	841	96.3	876	3	US-10-303-128-4	Sequence 4, Appli
26	841	96.3	876	5	US-10-735-419-4	Sequence 4, Appli
27	834.6	95.6	11474	3	US-09-495-406-1	Sequence 1, Appli
28	834.6	95.6	11474	3	US-09-816-028A-1	Sequence 1, Appli
29	834.6	95.6	11474	3	US-10-303-162-1	Sequence 1, Appli
30	834.6	95.6	11474	3	US-10-303-134-1	Sequence 1, Appli
31	834.6	95.6	11474	3	US-10-303-118-1	Sequence 1, Appli
32	834.6	95.6	11474	3	US-10-303-128-1	Sequence 1, Appli
33	834.6	95.6	11474	5	US-10-735-419-1	Sequence 1, Appli
34	833	95.4	876	3	US-09-495-406-2	Sequence 2, Appli
35	833	95.4	876	3	US-09-816-028A-2	Sequence 2, Appli
36	833	95.4	876	3	US-10-303-162-2	Sequence 2, Appli
37	833	95.4	876	3	US-10-303-134-2	Sequence 2, Appli
38	833	95.4	876	3	US-10-303-118-2	Sequence 2, Appli
39	833	95.4	876	3	US-10-303-128-2	Sequence 2, Appli
40	833	95.4	876	5	US-10-735-419-2	Sequence 2, Appli
41	817	93.6	876	3	US-09-495-406-6	Sequence 6, Appli
42	817	93.6	876	3	US-09-816-028A-6	Sequence 6, Appli
43	817	93.6	876	3	US-10-303-162-6	Sequence 6, Appli
44	817	93.6	876	3	US-10-303-134-6	Sequence 6, Appli
45	817	93.6	876	3	US-10-303-118-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-816-028A-13  
; Sequence 13, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-09-816-028A-13

Query Match		100.0%;	Score 873;	DB 3;	Length 873;
Best Local Similarity		100.0%;	Pred. No. 1.4e-158;		
Matches 873;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTC	CAAGG	60	
Db	1	ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTC	CAAGG	60	
Qy	61	CTACCAAAATGATTTTGTATTTAGATGATTAATCAATTTTATTTTGAAGATAAT	TACTAT	120	
Db	61	CTACCAAAATGATTTTGTATTTAGATGATTAATCAATTTTATTTTGAAGATAAT	TACTAT	120	
Qy	121	CTTGTAATAAATGCAAAACAGTGTTTACACCCCTAATTTCTTCTTGAGCAAT	TACTAC	180	
Db	121	CTTGTAATAAATGCAAAACAGTGTTTACACCCCTAATTTCTTCTTGAGCAAT	TACTAC	180	





```

RESULT 3
US-10-303-134-13
; Sequence 13, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-134-13

Query Match      100.0%; Score 873; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.4e-158;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTGCTGGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
DB 1 ATGAAAAAGTTATTTGCTGGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
QY 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTGAGATAATCTAT 120
DB 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTGAGATAATCTAT 120
QY 121 CTGTTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180
DB 121 CTGTTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180
QY 181 ACTTTAAAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTAAT 240
QY 241 TACACCAAGCTCATAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTCT 300
DB 241 TACACCAAGCTCATAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
QY 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420
DB 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420
QY 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600

```

```

DB 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAACTTACAAAAATAAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
DB 601 GAAAAAACTTACAAAAATAAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
QY 721 AAAGATATATCACTACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAAATTTT 780
DB 721 AAAGATATATCACTACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAAATTTT 780
QY 781 AAAAAAATAAAAAATAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
DB 781 AAAAAAATAAAAAATAAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873
DB 841 CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 4
US-10-303-118-13
; Sequence 13, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-118-13

Query Match      100.0%; Score 873; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.4e-158;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTGCTGGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
DB 1 ATGAAAAAGTTATTTGCTGGAAATGACCAAGTTTAAAGAAATGATTATTCAGG 60
QY 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTGAGATAATCTAT 120
DB 61 CTACCAATGATTTGATGTTATTTAGATGTAATCAATTTTATTTGAGATAATCTAT 120
QY 121 CTGTTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180
DB 121 CTGTTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATAC 180
QY 181 ACTTTAAAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATATGTTCTAAT 240
QY 241 TACACCAAGCTCATAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTCT 300
DB 241 TACACCAAGCTCATAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
QY 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420
DB 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTTGAGTAGCC 420
QY 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
DB 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTAGAAATTTCTA 600

```



US-10-735-419-13  
 ; Sequence I3, Application US/10735419  
 ; Patent No. 7026147  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000110US  
 ; CURRENT APPLICATION NUMBER: US/10/735,419  
 ; PRIOR FILING DATE: 2003-12-11  
 ; PRIOR APPLICATION NUMBER: US/09/816,028A  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 13  
 ; LENGTH: 873  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(873)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:36  
 US-10-735-419-13

Query Match 100.0%; Score 873; DB 5; Length 873;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-158;  
 Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGCTTTAAAGAAATTCATTATTCAGG 60  
 DB 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGCTTTAAAGAAATTCATTATTCAGG 60  
 QY 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATCTAT 120  
 DB 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATCTAT 120  
 QY 121 CTGGTAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180  
 DB 121 CTGGTAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180  
 QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAAGCTTAATATGTTCTTAAT 240  
 DB 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAAGCTTAATATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTGCT 300  
 DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTGCT 300  
 QY 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTAAA 360  
 DB 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTAAA 360  
 QY 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420  
 DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTC 480  
 DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTC 480  
 QY 481 TCTTATGCTTTTGTATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
 DB 481 TCTTATGCTTTTGTATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
 QY 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600

DB 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
 QY 601 GAAAAAATTTACAAAAATATAATATGCTTATGCTCTCTAATAGTCTTTTAGCAAAATTT 660  
 DB 601 GAAAAAATTTACAAAAATATAATATGCTTATGCTCTCTAATAGTCTTTTAGCAAAATTT 660  
 QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
 DB 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720  
 QY 721 AAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780  
 DB 721 AAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTT 780  
 QY 781 AAAAAAATAAAAATTAAGAAAAATGTTTATTTCAAGTTGATAAAGATCTATTAAAGATTA 840  
 DB 781 AAAAAAATAAAAATTAAGAAAAATGTTTATTTCAAGTTGATAAAGATCTATTAAAGATTA 840  
 QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
 DB 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 7

US-09-495-406-8  
 ; Sequence 8, Application US/09495406  
 ; Patent No. 6503744  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000110US  
 ; CURRENT APPLICATION NUMBER: US/09/495,406  
 ; CURRENT FILING DATE: 2000-01-31  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 US-09-495-406-8

Query Match 99.6%; Score 869.8; DB 3; Length 876;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-158;  
 Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGCTTTAAAGAAATTCATTATTCAGG 60  
 DB 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGCTTTAAAGAAATTCATTATTCAGG 60  
 QY 61 CTACCAATGATTTTGTATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATCTAT 120  
 DB 61 CTACCAATGATTTTGTATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATCTAT 120  
 QY 121 CTGGTAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180  
 DB 121 CTGGTAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTTGAGCAATCTAC 180  
 QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAAGCTTAATATGTTCTTAAT 240  
 DB 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAAGCTTAATATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCT 300

Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACGATATATTTTCCT 300  
Qy 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAAGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAAGCTTATTTTAAA 360  
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA 600  
Qy 601 GAAATACTTACAAATTAAGAAATATGCTTATCTTCTATAGCTTTTAAAGAAATTTT 660  
Db 601 GAAATACTTACAAATTAAGAAATATGCTTATCTTCTATAGCTTTTAAAGAAATTTT 660  
Qy 661 ATAGAACTAGCGCCCAATTTAAATTTCAATTTTATCATACAAAGAAATTAACACTACACT 720  
Db 661 ATAGAACTAGCGCCCAATTTAAATTTCAATTTTATCATACAAAGAAATTAACACTACACT 720  
Qy 721 AAGATATATCATACCTTCTAGTGGCTTTATGGAATTTTCAAAATAATTTTAAATTTT 780  
Db 721 AAGATATATCATACCTTCTAGTGGCTTTATGGAATTTTCAAAATAATTTTAAATTTT 780  
Qy 781 AAAAAATAAATAAAGAAATGTTTATTAAGTTGATATAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAATAAAGAAATGTTTATTAAGTTGATATAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

## RESULT 8

US-09-816-028A-8  
; Sequence 8, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816.028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-09-816-028A-8

Query Match 99.6%; Score 869.8; DB 3; Length 876;  
Best Local Similarity 99.8%; Pred. No. 5.8e-158;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60  
Db 1 ATGAAAAAGTTATTTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60  
Qy 61 CTACCAAAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 121 CTGCTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Db 121 CTGCTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATAGAGCCGAATTAATTTGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATAGAGCCGAATTAATTTGTTCTTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTTTCCT 300  
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTTTCCT 300  
Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAATGCTTTTAAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAATGCTTTTAAA 360  
Qy 361 TTTCACGAAATTTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Db 361 TTTCACGAAATTTTCAATCAAGAAATTTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTATCTTTCGGGAATTTTAT 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTTATCTTTCGGGAATTTTAT 480  
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA 600  
Db 541 GATCGCTCGCACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA 600  
Qy 601 GAAAAAATTTTCAAAATAAAGCTTATGCTTATGCTTCTTATGCTTTTACCAATTTT 660  
Db 601 GAAAAAATTTTCAAAATAAAGCTTATGCTTATGCTTCTTATGCTTTTACCAATTTT 660  
Qy 661 ATAGAACTAGCGCCCAATTTTAAATTTCAATTTTATCATACAAAGAAATTAACACTACACT 720  
Db 661 ATAGAACTAGCGCCCAATTTTAAATTTCAATTTTATCATACAAAGAAATTAACACTACACT 720  
Qy 721 AAGATATATCATACCTTCTAGTGGCTTTATGGAATTTTCAAAATAATTTTAAATTTT 780  
Db 721 AAGATATATCATACCTTCTAGTGGCTTTATGGAATTTTCAAAATAATTTTAAATTTT 780  
Qy 781 AAAAAATAAATAAAGAAATGTTTATTAAGTTGATATAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAATAAAGAAATGTTTATTAAGTTGATATAAGATCTATTAAAGATTA 840  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

## RESULT 9

US-10-303-162-8  
; Sequence 8, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada

;; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
;; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
;; FILE REFERENCE: 019633-000111US  
;; CURRENT APPLICATION NUMBER: US/10/303,162  
;; CURRENT FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US/09/816,028  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/118,213  
;; PRIOR FILING DATE: 1999-02-01  
;; PRIOR APPLICATION NUMBER: US 09/495,406  
;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 8  
;; LENGTH: 876  
;; TYPE: DNA  
;; ORGANISM: Campylobacter jejuni  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(876)  
;; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
;; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-162-8

Query Match 99.6%; Score 869.8; DB 3; Length 876;  
Best Local Similarity 99.8%; Pred. No. 5.8e-158;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGCAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTTACAGGCAATACTAC 180  
DB 121 CTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTTACAGGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATATGAGCCGAATTAATGTGTCTTAAT 240  
DB 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATATGAGCCGAATTAATGTGTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
QY 301 GATGCTCATTGGGCAAAATTTTAAATTTTATATCATACAAGAAAAATAACTACTACT 720

DB 661 ATGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACTACT 720  
QY 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
DB 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAGATTA 840  
DB 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAGATTA 840  
QY 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873  
DB 841 CCTAGTATATAAGCAATTTTCAAGGAAAA 873  
RESULT 10  
US-10-303-134-8  
;; Sequence 8, Application US/10303134  
;; Patent No. 6825019  
;; GENERAL INFORMATION:  
;; APPLICANT: Gilbert, Michel  
;; APPLICANT: Wakarchuk, Warren W.  
;; APPLICANT: National Research Council of Canada  
;; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
;; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
;; FILE REFERENCE: 019633-000111US  
;; CURRENT APPLICATION NUMBER: US/10/303,134  
;; CURRENT FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US/09/816,028  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/118,213  
;; PRIOR FILING DATE: 1999-02-01  
;; PRIOR APPLICATION NUMBER: US 09/495,406  
;; PRIOR FILING DATE: 2000-01-31  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 8  
;; LENGTH: 876  
;; TYPE: DNA  
;; ORGANISM: Campylobacter jejuni  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (1)..(876)  
;; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
;; OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-10-303-134-8

Query Match 99.6%; Score 869.8; DB 3; Length 876;  
Best Local Similarity 99.8%; Pred. No. 5.8e-158;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
DB 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTTCTTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTTCTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATATGAGCCGAATTAATGTGTCTTAAT 240  
DB 181 ACTTTAAACATTTTAAATCAAGAAATCAAGAAATATGAGCCGAATTAATGTGTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
QY 301 GATGCTCATTGGGCAAAATTTTAAATTTTATATCATACAAGAAAAATAACTACTACT 360

Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360  
Qy 361 TTTTCAAGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTCAGTAGCC 420  
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAT 540  
Qy 541 GATGCTCACACTATATCGGACATAGTAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGACATAGTAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Db 601 GAAAAAATTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Qy 721 AAAGATATACCTATACACTTTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 780  
Db 721 AAAGATATACCTATACACTTTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 780  
Qy 781 AAAAAATAAAATAAAGAAATGTTTATACAAAGTTGATATAAGATCTATTAAGATTA 840  
Db 781 AAAAAATAAAATAAAGAAATGTTTATACAAAGTTGATATAAGATCTATTAAGATTA 840  
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAA 873

## RESULT 11

US-10-303-118-8  
; Sequence 8, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni 0.19  
US-10-303-118-8

Query Match 99.6%; Score 869.8; DB 3; Length 876;

Best Local Similarity 99.8%; Pred. No. 5.8e-158;  
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTATTGCTGCAATGGACCAAGTTTAAAAAGAAATTCGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGCAATGGACCAAGTTTAAAAAGAAATTCGATTATTTCAAGG 60  
Qy 61 CTACCAATGATTTTGTATGATTTTATAGATGATTAATCAATTTTATTTTGAAGATAAAATCTAT 120  
Db 61 CTACCAATGATTTTGTATGATTTTATAGATGATTAATCAATTTTATTTTGAAGATAAAATCTAT 120  
Qy 121 CTTGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTGGAGCAATCTAC 180  
Db 121 CTTGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTGGAGCAATCTAC 180  
Qy 181 ACTTTAAAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATTAATGCTTCTAAAT 240  
Db 181 ACTTTAAAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATTAATGCTTCTAAAT 240  
Qy 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTTCGATTTATTTTCCCT 300  
Db 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACCTTTTTCGATTTATTTTCCCT 300  
Qy 301 GATGCTCATTTGGGATATGATTTTAAAACTTTAAAGAAATTTAAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAAACTTTAAAGAAATTTAAATGCTTATTTTAAA 360  
Qy 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTCTCAGGGGTCTATATGTCAGTAGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 541 GATGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
Db 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
Qy 601 GAAAAAATTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Db 601 GAAAAAATTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Qy 661 ATAGAACTAGCGCCAAATTTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTTAAATCAAAATTTTATCATACAGAAAAATAACTACACT 720  
Qy 721 AAAGATATACCTATACACTTTTAAATCAAAATTTTATGAGGAAATTTTCAAAATTAATTTT 780  
Db 721 AAAGATATACCTATACACTTTTAAATCAAAATTTTATGAGGAAATTTTCAAAATTAATTTT 780  
Qy 781 AAAAAATAAAATAAAGAAATGTTTATTAACAAGTTGATATAAGATCTATTAAGATTA 840  
Db 781 AAAAAATAAAATAAAGAAATGTTTATTAACAAGTTGATATAAGATCTATTAAGATTA 840  
Qy 841 CCTAGTATATAAGCAATTTTCAAGGAAA 873  
Db 841 CCTAGTATATAAGCAATTTTCAAGGAAA 873

## RESULT 12

US-10-303-128-8  
; Sequence 8, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of



```

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-128-8

Query Match          99.6%; Score 869.8; DB 3; Length 876;
Best Local Similarity 99.8%; Pred. No. 5.8e-158;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTGTAAAGAAATGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTGTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAATGATTTGATGATATTTAGATGTAATCAATTTATTTTGAAGATAATCTAT 120
DB 61 CTACCAATGATTTGATGATATTTAGATGTAATCAATTTATTTTGAAGATAATCTAT 120

QY 121 CTTCGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGGCAATCTAC 180
DB 121 CTTCGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGGCAATCTAC 180

QY 181 ACTTTAAACATTTAATCCAAAATCAAGAAATGAGCCGAATAATATTATGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAAATCAAGAAATGAGCCGAATAATATTATGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACCGATTATTTTCC 300
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACCGATTATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAATCTAAATTTTATCATACAGAAAAATACTACACT 720
DB 301 GATGCTCATTTGGGATATGATTTTAAATCTAAATTTTATCATACAGAAAAATACTACACT 720

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-735-419-8

Query Match          99.6%; Score 869.8; DB 5; Length 876;
Best Local Similarity 99.8%; Pred. No. 5.8e-158;
Matches 871; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTGTAAAGAAATGATTATTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTGTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAATGATTTGATGATATTTAGATGTAATCAATTTATTTTGAAGATAATCTAT 120
DB 61 CTACCAATGATTTGATGATATTTAGATGTAATCAATTTATTTTGAAGATAATCTAT 120

QY 121 CTTCGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGGCAATCTAC 180
DB 121 CTTCGTAATAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGGCAATCTAC 180

QY 181 ACTTTAAACATTTAATCCAAAATCAAGAAATGAGCCGAATAATATTATGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCCAAAATCAAGAAATGAGCCGAATAATATTATGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCC 300
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360

QY 361 TTTCAACGAAATTTATTTCATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
DB 361 TTTCAACGAAATTTATTTCATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCATTATGCTTTTGTATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
DB 481 TCATTATGCTTTTGTATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATGCTCATCTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
DB 541 GATGCTCATCTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATCTTACAAAATATAATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660
DB 601 GAAAAAATCTTACAAAATATAATATGCTTATGCTTCTTAATAGTCTTTTACGAAATTTT 660

QY 661 ATAGAACTAGCGCAAAATTTAAATCTAAATTTTATCATACAGAAAAATACTACACT 720
DB 661 ATAGAACTAGCGCAAAATTTAAATCTAAATTTTATCATACAGAAAAATACTACACT 720

```



Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAGAAATTAATGCTTATTTTAA 360  
 Qy 361 TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGTCTATATGTGCGAGTAGCC 420  
 Db 361 TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGTCTATATGTGCGAGTAGCC 420  
 Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Qy 541 GATGCTCACACTATATCGGCATAGTAAAGAAATACAGATATAAGAGCTTTTGAATTTCTA 600  
 Db 541 GATGCTCGCACATATATCGGCATAGTAAAGAAATACAGATATAAGAGCTTTTGAATTTCTA 600  
 Qy 601 GAAAAAATTTTACAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 720  
 Db 601 GAAAAAATTTTACAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 720  
 Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 720  
 Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAAATTTTCAAAATTAATTTT 780  
 Db 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAAATTTTCAAAATTAATTTT 780  
 Qy 781 AAAAAATTAAGAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 840  
 Db 781 AAAAAATTAAGAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 840  
 Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
 Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

RESULT 14

US-09-816-028A-11  
 ; Sequence 11, Application US/09816028A  
 ; Patent No. 6699705  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/09/816,028A  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 11  
 ; LENGTH: 873  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(873)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:4  
 US-09-816-028A-11

Query Match 99.3%; Score 866.6; DB 3; Length 873;  
 Best Local Similarity 99.5%; Pred. No. 2.4e-157;  
 Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTATCAAGG 60  
 Db 1 ATGAAAAAGTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTATCAAGG 60  
 Qy 61 CTACCAAAATGATTTTATGATGATTTTATAGATCAATTTTATTTTGAAGATAAATACTAT 120  
 Db 61 CTACCAAAATGATTTTATGATGATTTTATAGATCAATTTTATTTTGAAGATAAATACTAT 120  
 Qy 121 CTGTGTAATAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
 Db 121 CTGTGTAATAATGCAAAACAGTGTTTTACACCCCTGGTTTCTTCTTTGAGCAATACTAC 180  
 Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATATATGCTTCTAAT 240  
 Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAATATATGCTTCTAAT 240  
 Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACCGATTAATTTCCCT 300  
 Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACCGATTAATTTCCCT 300  
 Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360  
 Qy 361 TTTCAGAAATTTATTTCAATCAAGAAATTACCTCAGGGTCTATATGTGCGAGTAGCC 420  
 Db 361 TTTCAGAAATTTATTTCAATCAAGAAATTACCTCAGGGTCTATATGTGCGAGTAGCC 420  
 Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Qy 541 GATGCTCACACTATATCGGCATAGTAAAGAAATACAGATATAAGAGCTTTTGAATTTCTA 600  
 Db 541 GATGCTCACACTATATCGGCATAGTAAAGAAATACAGATATAAGAGCTTTTGAATTTCTA 600  
 Qy 601 GAAAAAATTTTACAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 660  
 Db 601 GAAAAAATTTTACAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 660  
 Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 720  
 Qy 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAAATTTTCAAAATTAATTTT 780  
 Db 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAAATTTTCAAAATTAATTTT 780  
 Qy 781 AAAAAATTAAGAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 840  
 Db 781 AAAAAATTAAGAAATTAAGAAATTTTAAATTTTATCATACAAAGAAATTAAGACT 840  
 Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873  
 Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

RESULT 15

US-10-303-162-11  
 ; Sequence 11, Application US/10303162  
 ; Patent No. 6723545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/303,162

```

; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US/09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-162-11

Query Match          99.3%; Score 866.6; DB 3; Length 873;
Best Local Similarity 99.5%; Pred. No. 2.4e-157;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTCGATTATTCAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTCGATTATTCAGG 60

QY 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
DB 61 CTACCAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTTGTTAAAAATGCAAAACAGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
DB 121 CTTGTTAAAAATGCAAAACAGTGTTTACACCCCTGTTCTTCTTGAGCAATACTAC 180

QY 181 ACTTTAAAAACATTTAATCAAAATCAAGATATGAGCGAATCAATTAATGTTCTAAT 240
DB 181 ACTTTAAAAACATTTAATCAAAATCAAGATATGAGCGAATCAATTAATGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTAAAGAAATTTAAAGAAATTTAAAGAAATTTTAA 300
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAA 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGAAATTTAAAGAAATTTTAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGAAATTTAAAGAAATTTTAA 360

QY 361 TTTACGAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420
DB 361 TTTACGAAATTTTATTTCAATCAAAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATCAAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATCAAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAA 540
DB 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAA 540

QY 541 GATCGCTCACAATATATCGGATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
DB 541 GATCGCTCACAATATATCGGATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTCAAAATTAATACTATATGCTTATGTCCTAATAGTCTTTTAGCAATTTT 660
DB 601 GAAAAAATTTCAAAATTAATACTATATGCTTATGTCCTAATAGTCTTTTAGCAATTTT 660

QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAGAAATTAATACTACT 720
DB 661 ATAGAACTAGCGCCAAATTTTAAATTTTAAATTTTATCATACAGAAATTAATACTACT 720

QY 721 AAAGATATCTATACCTTCTTAGTGAGGCTTTATGGAAATTTTCAAAAAATATTAATTTT 780
```

```

|||||
721 AAAGATATCTATACCTTCTTAGTGAGGCTTTATGGAAATTTTCAAAAAATATTAATTTT 780
QY
781 AAAAAAATAAAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
|||||
781 AAAAAAATAAAAAATTAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
|||||
841 CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873
DB
841 CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873
|||||
```

Search completed: July 19, 2006, 15:09:21  
Job time : 138.177 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: July 19, 2006, 15:09:43 ; Search time 996.192 Seconds  
(without alignments)  
10768.108 Million cell updates/sec  
Title: US-10-734-719-13  
Perfect score: 873  
Sequence: 1 atgaaaaaagtattattgc.....agcattatttcaaggaaaaa 873  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 18892170 seqs, 6143817638 residues  
Total number of hits satisfying chosen parameters: 37784340  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.Main.\*  
1: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
3: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
4: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
5: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
6: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
7: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
8: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
9: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
10: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
11: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
12: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
13: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
14: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US11B\_PUBCOMB.seq.\*  
15: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US11C\_PUBCOMB.seq.\*  
16: /EMC\_Celerra\_SID33/ptodata/2/pubpna/US11D\_PUBCOMB.seq.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Query		Match		Length		DB		ID		Description	
Result No.	Score	Query Match	Score	Match	Length	DB	Score	Match	Length	DB	Score	Match	Length	ID	Description
1	873	100.0	873	3	US-09-816-028A-13	Sequence 13, Appl									
2	873	100.0	873	7	US-10-303-161-13	Sequence 13, Appl									
3	873	100.0	873	7	US-10-303-118-13	Sequence 13, Appl									
4	873	100.0	873	7	US-10-303-128-13	Sequence 13, Appl									
5	873	100.0	873	7	US-10-303-134-13	Sequence 13, Appl									
6	873	100.0	873	7	US-10-303-162-13	Sequence 13, Appl									
7	873	100.0	873	9	US-10-735-419-13	Sequence 13, Appl									
8	873	100.0	873	9	US-10-820-536-13	Sequence 13, Appl									
9	873	100.0	873	9	US-10-845-408-13	Sequence 13, Appl									
10	873	100.0	873	9	US-10-845-412-13	Sequence 13, Appl									
11	873	100.0	873	9	US-10-846-219-13	Sequence 13, Appl									
12	873	100.0	873	9	US-10-821-604-13	Sequence 13, Appl									
13	873	100.0	873	9	US-10-847-983-13	Sequence 13, Appl									
14	873	100.0	873	9	US-10-821-573-13	Sequence 13, Appl									
15	873	100.0	873	9	US-10-850-807-13	Sequence 13, Appl									
16	873	100.0	873	9	US-10-850-125-13	Sequence 13, Appl									
17	873	100.0	873	9	US-10-830-825-13	Sequence 13, Appl									

18	873	100.0	873	10	US-10-962-334-13	Sequence 13, Appl
19	873	100.0	873	10	US-10-830-997-13	Sequence 13, Appl
20	873	100.0	873	10	US-10-962-235-13	Sequence 13, Appl
21	873	100.0	873	10	US-10-961-882-13	Sequence 13, Appl
22	869.8	99.6	876	3	US-09-816-028A-8	Sequence 8, Appl
23	869.8	99.6	876	7	US-10-303-161-8	Sequence 8, Appl
24	869.8	99.6	876	7	US-10-303-118-8	Sequence 8, Appl
25	869.8	99.6	876	7	US-10-303-128-8	Sequence 8, Appl
26	869.8	99.6	876	7	US-10-303-134-8	Sequence 8, Appl
27	869.8	99.6	876	7	US-10-303-162-8	Sequence 8, Appl
28	869.8	99.6	876	9	US-10-735-419-8	Sequence 8, Appl
29	869.8	99.6	876	9	US-10-820-536-8	Sequence 8, Appl
30	869.8	99.6	876	9	US-10-845-408-8	Sequence 8, Appl
31	869.8	99.6	876	9	US-10-845-412-8	Sequence 8, Appl
32	869.8	99.6	876	9	US-10-846-219-8	Sequence 8, Appl
33	869.8	99.6	876	9	US-10-821-604-8	Sequence 8, Appl
34	869.8	99.6	876	9	US-10-847-983-8	Sequence 8, Appl
35	869.8	99.6	876	9	US-10-821-573-8	Sequence 8, Appl
36	869.8	99.6	876	9	US-10-850-807-8	Sequence 8, Appl
37	869.8	99.6	876	9	US-10-850-125-8	Sequence 8, Appl
38	869.8	99.6	876	9	US-10-830-825-8	Sequence 8, Appl
39	869.8	99.6	876	10	US-10-962-334-8	Sequence 8, Appl
40	869.8	99.6	876	10	US-10-830-997-8	Sequence 8, Appl
41	869.8	99.6	876	10	US-10-962-235-8	Sequence 8, Appl
42	869.8	99.6	876	10	US-10-961-882-8	Sequence 8, Appl
43	866.6	99.3	873	3	US-09-816-028A-11	Sequence 11, Appl
44	866.6	99.3	873	7	US-10-303-161-11	Sequence 11, Appl
45	866.6	99.3	873	7	US-10-303-118-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-09-816-028A-13  
; Sequence 13, Application US/09816028A  
; Patent No. US20020042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816, 028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-09-816-028A-13

Query Match 100.0%; Score 873; DB 3; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Qy 61 CTACCAATGATTTTGTATTTAGATGATCAATTTTATTTTGAAGATAAATCTAT 120

```
Db 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGAATATGAGACCGAAGCTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAATATGAGACCGAAGCTAAATATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Qy 361 TTTTACCAAAATTTTATTTCAATCAAGAATTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420
Db 361 TTTTACCAAAATTTTATTTCAATCAAGAATTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATGCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATGCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACATATTTGCTTATGCTCTAATAGCTTTTACGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACATATTTGCTTATGCTCTAATAGCTTTTACGCAAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 721 AAAGATATCTATCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 780
Db 721 AAAGATATCTATCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 780
Qy 781 AAAAAATAAATAAATAAGAAATTTTATTTATACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAATAAATAAGAAATTTTATTTATACAGTTGATAAAGATCTATTAAAGATTA 840
```

## RESULT 2

```
US-10-303-161-13
; Sequence 13, Application US/10303161
; Publication No. US20030148459A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
```

```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-161-13
```

```
Query Match 100.0%; Score 873; DB 7; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.2e-136;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTTATTTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTTATTTCAAGATAAATACTAT 120
Db 1 ATGAAAAAGTTTATTTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTTATTTCAAGATAAATACTAT 120
Qy 61 CTACCAATGATTTTGTATTTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATTTTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 121 CTTCGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db 121 CTTCGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGAATATGAGACCGAAGCTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAATATGAGACCGAAGCTAAATATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTTATTTTCCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGCTATATATGTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATGCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATGCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACATATTTGCTTATGCTCTAATAGCTTTTACGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACATATTTGCTTATGCTCTAATAGCTTTTACGCAAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 721 AAAGATATCTATCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 780
Db 721 AAAGATATCTATCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 780
Qy 781 AAAAAATAAATAAATAAGAAATTTTATTTATACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAATAAATAAGAAATTTTATTTATACAGTTGATAAAGATCTATTAAAGATTA 840
```

QY 841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873  
Db |||||  
841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873  
|||  
RESULT 3  
US-10-303-118-13  
; Sequence 13, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-118-13

Query Match 100.0%; Score 873; DB 7; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
Db |||||  
1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
|||  
QY 61 CTACCAATGATTTTGATGTTTATGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db |||||  
61 CTACCAATGATTTTGATGTTTATGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
|||  
QY 121 CTGTTAAAAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
Db |||||  
121 CTGTTAAAAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
|||  
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATTAATGTTGTTCTAAT 240  
Db |||||  
181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATTAATGTTGTTCTAAT 240  
|||  
QY 241 TACACCAAGCTCATAGAAATGAAATTTTGTAAAACTTTTACAGTATTTTCT 300  
Db |||||  
241 TACACCAAGCTCATAGAAATGAAATTTTGTAAAACTTTTACAGTATTTTCT 300  
|||  
QY 301 GATGCTCATTTGGGATGATGTTTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
Db |||||  
301 GATGCTCATTTGGGATGATGTTTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360  
|||  
QY 361 TTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db |||||  
361 TTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420  
|||  
QY 421 ATAGCCCTAGATACAAAGAAATTTAATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db |||||  
421 ATAGCCCTAGATACAAAGAAATTTAATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
|||

QY 481 TCTTATGCTTTTGATATACCAAAACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db |||||  
481 TCTTATGCTTTTGATATACCAAAACAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
|||  
QY 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAAAGTTTCTA 600  
Db |||||  
541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTAAAGTTTCTA 600  
|||  
QY 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAAAGTTTCTA 660  
Db |||||  
601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAAAGTTTCTA 660  
|||  
QY 661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTACT 720  
Db |||||  
661 ATAGAACTAGCGCCAAATTTTAAATTTTATCATACAGAAAAATAAATACTACT 720  
|||  
QY 721 AAGATATATCTATACCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATAATTAATTT 780  
Db |||||  
721 AAGATATATCTATACCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATAATTAATTT 780  
|||  
QY 781 AAAAAATAAAAAATAAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTTAAAGATT 840  
Db |||||  
781 AAAAAATAAAAAATAAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTTAAAGATT 840  
|||  
QY 841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873  
Db |||||  
841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873  
|||

RESULT 4  
US-10-303-128-13  
; Sequence 13, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-128-13

Query Match 100.0%; Score 873; DB 7; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
Db |||||  
1 ATGAAAAAGTTATTCTGCGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60  
|||  
QY 61 CTACCAATGATTTTGATGTTTATGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db |||||  
61 CTACCAATGATTTTGATGTTTATGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
|||

```
Qy 121 CTTGGTAAAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATAC 180
Db 121 CTTGGTAAAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTCTTTGAGCAATAC 180
Qy 181 ACTTTAAACACTTAATTCACAAATCAAGATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACACTTAATTCACAAATCAAGATATGAGCCGAACTAATTTATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATATTATTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATATTATTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Qy 361 TTTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Db 361 TTTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACGAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACGAAATTT 660
Qy 661 ATAGAATACCTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATTAAC 720
Db 661 ATAGAATACCTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATTAAC 720
Qy 721 AAAGATATCTATACACTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATCTATACACTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATTAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
```

RESULT 5  
US-10-303-134-13  
; Sequence 13, Application US/10303134  
; Publication No. US20030157657A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR FILING DATE: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31

```
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3(alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-134-13
```

Query Match 100.0%; Score 873; DB 7; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGAAAAAGTTATTTGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTTATTCAGG 60
Db 1 ATGAAAAAGTTATTTGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTTATTCAGG 60
Qy 61 CTACCAAAATGATTTTATGATCTATTTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTATGATCTATTTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120
Qy 121 CTTGTAAAAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Db 121 CTTGTAAAAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTTTGAGCAATACTAC 180
Qy 181 ACTTTAAACATTTAATTCACAAATCAAGAAATGAGACCGAACTAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATTCACAAATCAAGAAATGAGACCGAACTAATTTATGTGTTCTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAATCTTTTACGATATTATTTCC 300
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAATCTTTTACGATATTATTTCC 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTTAAAGAAATTTAATGCTTTATTTAAA 360
Qy 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACGAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTTATGCTCTAATAGTCTTTTACGAAATTT 660
Qy 661 ATAGAATACCTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACTAC 720
Db 661 ATAGAATACCTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACTAC 720
Qy 721 AAAGATATCTATACACTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATCTATACACTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATTAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
```

QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873  
|||||  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

## RESULT 6

US-10-303-162-13  
; Sequence 13, Application US/10303162  
; Publication No. US20030157658A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-162-13

Query Match 100.0%; Score 873; DB 7; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTCGTCGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG 60  
Db 1 ATGAAAAAGTTATTATTCGTCGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG 60  
QY 61 CTACCAAAATGATTTTGATGTTATTTAGATGTTATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGATGTTATTTAGATGTTATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAAGTAAATATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAAGTAAATATGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAAAGAAATTTTGAAGAACTTTTACGATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAAAGAAATTTTGAAGAACTTTTACGATTTTCT 300  
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360  
QY 361 TTTACGAAATTTTAAATCAAGAAATACCTCAGGGGTCTATATGTTGACGATGCC 420  
Db 361 TTTACGAAATTTTAAATCAAGAAATACCTCAGGGGTCTATGTTGTTGACGATGCC 420  
QY 421 ATAGCCCTAGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
|||||  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600  
|||||  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600  
QY 601 GAAAAAACTTTACAAAAATAAACTATATTGCTTATGCTCTCTAATAGCTTTTGAATTTT 660  
|||||  
Db 601 GAAAAAACTTTACAAAAATAAACTATATTGCTTATGCTCTCTAATAGCTTTTGAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACT 720  
|||||  
Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACT 720  
QY 721 AAAGATATATCTCATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTT 780  
|||||  
Db 721 AAAGATATATCTCATACCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTT 780  
QY 781 AAAAAAATAAAAAATTTAAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
|||||  
Db 781 AAAAAAATAAAAAATTTAAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873  
|||||  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

## RESULT 7

US-10-735-419-13  
; Sequence 13, Application US/10735419  
; Publication No. US20040180406A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-735-419-13

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCGTCGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG 60  
Db 1 ATGAAAAAGTTATTATTCGTCGAATGGACCAAGTTTAAAGAAATTCATTATTCAGG 60  
QY 61 CTACCAAAATGATTTTGATGTTATTTAGATGTTATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGATGTTATTTAGATGTTATCAATTTTATTTGAAGATAAATACTAT 120

```
QY 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACATAC 180
Db 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACATAC 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCGGAACTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCGGAACTAAATATGTGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATATTTTCT 300
QY 301 GATGCTCATTTGGGATGATTTTAAAAAACAATTTAAAGAAATTTAAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATGATTTTAAAAAACAATTTAAAGAAATTTAAATGCTTATTTTAAA 360
QY 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAAAATCTTTTAAATGCTTATTTTAAA 540
Db 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAAAATCTTTTAAATGCTTATTTTAAA 540
QY 541 GATGCTCAGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
Db 541 GATGCTCAGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
QY 601 GAAAAAATCTTCAAAATAAAAATATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 601 GAAAAAATCTTCAAAATAAAAATATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
QY 721 AAAGATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 780
Db 721 AAAGATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 780
QY 781 AAAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATTAAGATTA 840
Db 781 AAAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATTAAGATTA 840
```

## RESULT 8

```
US-10-820-536-13
; Sequence 13, Application US/10820536
; Publication No. US20040203103A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
```

```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0:36
; US-10-820-536-13
```

```
Query Match 100.0%; Score 873; DB 9; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.2e-136;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAAGTTTATTTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTTATTTCAAGATTAATCTAT 120
Db 1 ATGAAAAAGTTTATTTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTTATTTCAAGATTAATCTAT 120
QY 61 CTACCAATGATTTTGTATGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACATAC 180
Db 121 CTTGGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTCTTTGAGCAATACATAC 180
QY 181 ACTTTAAAAATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAAATTTATGTGTTCTAAT 240
Db 181 ACTTTAAAAATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAAATTTATGTGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATTTATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATTTATTTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAAAAACAATTTTAAAGAAATTTTAAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAAAAACAATTTTAAAGAAATTTTAAATGCTTATTTTAAA 360
QY 361 TTTTCAAGAAATTTTAAATCCAAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTAAATCCAAATCAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAAAATCTTTTAAATGCTTATTTTAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAAAATCTTTTAAATGCTTATTTTAAAAT 540
QY 541 GATGCTCAGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
Db 541 GATGCTCAGACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTAAATTTCTA 600
QY 601 GAAAAAATCTTCAAAATAAAAATATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 601 GAAAAAATCTTCAAAATAAAAATATATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 720
QY 721 AAAGATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 780
Db 721 AAAGATATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 780
QY 781 AAAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATTAAGATTA 840
Db 781 AAAAAAATAAAAATTTAAAGAAATGTTTATTAACAAGTTGATTAAGATTA 840
```



QY	841	CCTAGTGATATAAGCAATTAATTTCAAAGGAAA	873
Db	841	CCTAGTGATATAAGCAATTAATTTCAAAGGAAA	873
RESULT 9			
US-10-845-408-13			
; Sequence 13, Application US/10845408			
; Publication No. US20040203112A1			
; GENERAL INFORMATION:			
; APPLICANT: Gilbert, Michel			
; APPLICANT: Wakarchuk, Warren W.			
; APPLICANT: National Research Council of Canada			
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of			
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics			
; FILE REFERENCE: 019633-000111US			
; CURRENT APPLICATION NUMBER: US/10/845,408			
; PRIOR FILING DATE: 2004-05-12			
; PRIOR APPLICATION NUMBER: US/09/816,028			
; PRIOR FILING DATE: 2001-03-21			
; PRIOR APPLICATION NUMBER: US 60/118,213			
; PRIOR FILING DATE: 1999-02-01			
; PRIOR APPLICATION NUMBER: US 09/495,406			
; PRIOR FILING DATE: 2000-01-31			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 13			
; LENGTH: 873			
; TYPE: DNA			
; ORGANISM: Campylobacter jejuni			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(873)			
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II			
; OTHER INFORMATION: (CstII) from C. jejuni O:36			
US-10-845-408-13			
Query Match 100.0%; Score 873; DB 9; Length 873;			
Best Local Similarity 100.0%; Pred. No. 2.2e-136;			
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGAAAAAGTTATTCTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60
Db	1	ATGAAAAAGTTATTCTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60
QY	61	CTACCAATGATTTTGGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120
Db	61	CTACCAATGATTTTGGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120
QY	121	CTTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACAC	180
Db	121	CTTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACAC	180
QY	181	ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAAATATGTTGTTCTAAT	240
Db	181	ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAAATATGTTGTTCTAAT	240
QY	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTTCAGTATTCTTCT	300
Db	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTTCAGTATTCTTCT	300
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA	360
Db	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA	360
QY	361	TTTCAGAAATTTATTCAATCAAGAAATACCTCAGGGGTCTATPATGTGTCAGTAGCC	420
Db	361	TTTCAGAAATTTATTCAATCAAGAAATACCTCAGGGGTCTATPATGTGTCAGTAGCC	420
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA	480
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA	480

QY	481	TCCTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
Db	481	TCCTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
QY	541	GATCGCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA	600
Db	541	GATCGCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTTAAAGATTTCTA	600
QY	601	GAAAAAATTTACAAAAATAAATACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTT	660
Db	601	GAAAAAATTTACAAAAATAAATACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTT	660
QY	661	ATAGAACTAGCGCAAAATTTAAATTTAAATTTATCATACAAGAAAAATAAATACTACT	720
Db	661	ATAGAACTAGCGCAAAATTTAAATTTAAATTTATCATACAAGAAAAATAAATACTACT	720
QY	721	AAAGATATATCTACATCTCTAGTGGCTTATGGAATTTTCAAAAAATATTAATTTT	780
Db	721	AAAGATATATCTACATCTCTAGTGGCTTATGGAATTTTCAAAAAATATTAATTTT	780
QY	781	AAAAAAATAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA	840
Db	781	AAAAAAATAAAAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA	840
QY	841	CCTAGTGATATAAGCAATTAATTTCAAAGGAAA	873
Db	841	CCTAGTGATATAAGCAATTAATTTCAAAGGAAA	873
RESULT 10			
US-10-845-412-13			
; Sequence 13, Application US/10845412			
; Publication No. US20040203113A1			
; GENERAL INFORMATION:			
; APPLICANT: Gilbert, Michel			
; APPLICANT: Wakarchuk, Warren W.			
; APPLICANT: National Research Council of Canada			
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of			
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics			
; FILE REFERENCE: 019633-000111US			
; CURRENT APPLICATION NUMBER: US/10/845,412			
; PRIOR FILING DATE: 2004-05-12			
; PRIOR APPLICATION NUMBER: US/10/303,128			
; PRIOR FILING DATE: 2002-11-21			
; PRIOR APPLICATION NUMBER: US/09/816,028			
; PRIOR FILING DATE: 2001-03-21			
; PRIOR APPLICATION NUMBER: US 60/118,213			
; PRIOR FILING DATE: 1999-02-01			
; PRIOR APPLICATION NUMBER: US 09/495,406			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 13			
; LENGTH: 873			
; TYPE: DNA			
; ORGANISM: Campylobacter jejuni			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(873)			
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II			
; OTHER INFORMATION: (CstII) from C. jejuni O:36			
US-10-845-412-13			
Query Match 100.0%; Score 873; DB 9; Length 873;			
Best Local Similarity 100.0%; Pred. No. 2.2e-136;			
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGAAAAAGTTATTCTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60
Db	1	ATGAAAAAGTTATTCTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60
QY	61	CTACCAATGATTTTGGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120
Db	61	CTACCAATGATTTTGGTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120
QY	121	CTTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACAC	180
Db	121	CTTGTAATAAATGCAAAACAGTGTTTTACACCCCTAATTTCTTTGAGCAATACAC	180
QY	181	ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAAATATGTTGTTCTAAT	240
Db	181	ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAAATATGTTGTTCTAAT	240
QY	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTTCAGTATTCTTCT	300
Db	241	TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTTCAGTATTCTTCT	300
QY	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA	360
Db	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATAA	360
QY	361	TTTCAGAAATTTATTCAATCAAGAAATACCTCAGGGGTCTATPATGTGTCAGTAGCC	420
Db	361	TTTCAGAAATTTATTCAATCAAGAAATACCTCAGGGGTCTATPATGTGTCAGTAGCC	420
QY	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA	480
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA	480

Db 61 CTACCAATGATTTTGGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Qy 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 240  
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATATTTTCT 300  
Qy 301 GATGCTCATTTGGGATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Qy 361 TTTCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 420  
Db 361 TTTCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 420  
Qy 421 ATAGCCCTAGCATCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 480  
Db 421 ATAGCCCTAGCATCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 480  
Qy 481 TCTTATGCTTTGATACCAACAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 540  
Db 481 TCTTATGCTTTGATACCAACAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 540  
Qy 541 GATCGCTCACAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 600  
Db 541 GATCGCTCACAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 600  
Qy 601 GAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 660  
Db 601 GAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 660  
Qy 661 ATAGAACTTAGCCCAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 720  
Db 661 ATAGAACTTAGCCCAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 720  
Qy 721 AAAGATATCTATACATCTTCTAGTGGGCTTATGAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 780  
Db 721 AAAGATATCTATACATCTTCTAGTGGGCTTATGAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 780  
Qy 781 AAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 840  
Db 781 AAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 840  
Qy 841 CCTAGTATATAAGCAATTTTCAAAAGGAAA 873  
Db 841 CCTAGTATATAAGCAATTTTCAAAAGGAAA 873

## RESULT 11

US-10-846-219-13  
; Sequence 13, Application US/10846219  
; Publication No. US20040219638A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/846,219  
; PRIOR FILING DATE: 2004-05-14  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2004-05-14  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; NAME/KEY: CDS  
; LOCATION: (1) .. (873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-846-219-13

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136; Indels 0; Gaps 0;  
Matches 873; Conservative 0; Mismatches 0;  
Qy 1 ATGAAAAAGTTTATTTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTTATTTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTATTTCAAGG 60  
Qy 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGTATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 121 CTTGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Db 121 CTTGTAAAAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180  
Qy 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 240  
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATATTTTCT 300  
Qy 301 GATGCTCATTTGGGATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Db 301 GATGCTCATTTGGGATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360  
Qy 361 TTTCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 420  
Db 361 TTTCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 420  
Qy 421 ATAGCCCTAGCATCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 480  
Db 421 ATAGCCCTAGCATCAAGAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 480  
Qy 481 TCTTATGCTTTGATACCAACAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 540  
Db 481 TCTTATGCTTTGATACCAACAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 540  
Qy 541 GATCGCTCACAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 600  
Db 541 GATCGCTCACAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 600  
Qy 601 GAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 660  
Db 601 GAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 660  
Qy 661 ATAGAACTTAGCCCAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 720  
Db 661 ATAGAACTTAGCCCAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 720  
Qy 721 AAAGATATCTATACATCTTCTAGTGGGCTTATGAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 780  
Db 721 AAAGATATCTATACATCTTCTAGTGGGCTTATGAGAAATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 780  
Qy 781 AAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 840  
Db 781 AAAAAAATTTAATCAAAATCAAGATATGAGACCGAATTAATTTATGTTCTAAT 840

Db 781 AAAAAAAAAAATAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTATTTCACAAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTATTTCACAAAGGAAAA 873

RESULT 12  
US-10-821-604-13  
; Sequence 13, Application US/10821604  
; Publication No. US20040229263A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-821-604-13

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATCTCGAAATGACCAAGTTTAAAGAAATTCGATTATTCACAGG 60  
Db 1 ATGAAAAAGTTATTATCTCGAAATGACCAAGTTTAAAGAAATTCGATTATTCACAGG 60

QY 61 CTACCAAAATGATTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTACCAAAATGATTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120

QY 121 CTGTGTAATAAATGCAAAACAGTGTGTTACACCCCTAAATTTCTTTGAGCAAACTACTAC 180  
Db 121 CTGTGTAATAAATGCAAAACAGTGTGTTACACCCCTAAATTTCTTTGAGCAAACTACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATGTTGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCACTAATGTTGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTTCT 300

QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAA 360

QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATCGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATCGGTCA 480

QY 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAATTTACAAAAATAAAACTATATGCTTATGCTTCTAATAGTCTTTTAGCAATTTT 660  
Db 601 GAAAAAATTTACAAAAATAAAACTATATGCTTATGCTTCTAATAGTCTTTTAGCAATTTT 660

QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAAATAACTACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAAATAACTACT 720

QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTTATGAAAAATTTTCAAAAAATTTTAT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTTATGAAAAATTTTCAAAAAATTTTAT 780

QY 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAGCAATTATTTCACAAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTATTTCACAAAGGAAAA 873

RESULT 13  
US-10-847-983-13  
; Sequence 13, Application US/10847983  
; Publication No. US20040229272A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/847,983  
; CURRENT FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-847-983-13

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTTGCTGAAATGACCAAGTTTAAAGAAATTTAGATTATTCAGG 60  
Db 1 ATGAAAAAGTTATTATTTGCTGAAATGACCAAGTTTAAAGAAATTTAGATTATTCAGG 60

```
QY 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT 120
Db 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT 120
QY 121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACAT 180
Db 121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACAT 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAATCAATTAATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAATCAATTAATGTTCTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTTTCT 300
Db 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTTTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTTTAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTTTAA 360
QY 361 TTTCAAGAAATTTAATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
Db 361 TTTCAAGAAATTTAATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGTCA 480
QY 481 TCTTATGCTTTTGAACAAATTTAAATCTTTTAAACAACTTAAAGAAATTTAAAGCTTTTAA 540
Db 481 TCTTATGCTTTTGAACAAATTTAAATCTTTTAAACAACTTAAAGAAATTTAAAGCTTTTAA 540
QY 541 GATCGCTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600
Db 541 GATCGCTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTAATAAATAAATTTAATTTCAATTTTATCATACAGAAAAATTAACACT 720
Db 601 GAAAAAATTTAATAAATAAATTTAATTTCAATTTTATCATACAGAAAAATTAACACT 720
QY 721 AAAGATATCTCATACCTTCTAGTGAGGCTTATGCGGAATTTTCAAAAAATTAATTTT 780
Db 721 AAAGATATCTCATACCTTCTAGTGAGGCTTATGCGGAATTTTCAAAAAATTAATTTT 780
QY 781 AAAAAATTTAATAAATAAATTTTATTTACAAATTTTCAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTTAATAAATAAATTTTATTTACAAATTTTCAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTATTAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTATTAAGCATTTTCAAGGAAAA 873
```

RESULT 14

US-10-821-573-13  
; Sequence 13, Application US/10821573  
; Publication No. US20040229313A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,573  
; PRIOR FILING DATE: 2004-04-08  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028

PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 13  
LENGTH: 873  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(873)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-821-573-13

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTATTTCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60  
Db 1 ATGAAAAAGTTATTTATTTCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT 120  
Db 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACAT 120  
QY 121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACAT 180  
Db 121 CTGGTAAAAATGCAAAACAGTGTCTTACACCCCTAAATTTCTTCTTGAGCAATACAT 180  
QY 181 ACTTTAAACATTTAATTTCAATTTCAAGAATTTAGACCCGAATTAATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATTTCAATTTCAAGAATTTAGACCCGAATTAATGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAATCTTTTACGATTTTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
QY 361 TTTCAAGAAATTTAATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTCAAGAAATTTAATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGTCA 480  
QY 481 TCTTATGCTTTTGAACAAATTTAAATCTTTTAAACAACTAGCCCTGATTTTAA 540  
Db 481 TCTTATGCTTTTGAACAAATTTAAATCTTTTAAACAACTAGCCCTGATTTTAA 540  
QY 541 GATCGCTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600  
Db 541 GATCGCTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAAATTTCTA 600  
QY 601 GAAAAAATTTAATAAATAAATTTAATTTCAATTTTATCATACAGAAAAATTAACACT 720  
Db 601 GAAAAAATTTAATAAATAAATTTAATTTCAATTTTATCATACAGAAAAATTAACACT 720  
QY 721 AAAGATATCTCATACCTTCTAGTGAGGCTTATGCGGAATTTTCAAAAAATTAATTTT 780  
Db 721 AAAGATATCTCATACCTTCTAGTGAGGCTTATGCGGAATTTTCAAAAAATTAATTTT 780

QY 781 AAAAAAATAAATAAAGAAATGTTATTACAAGTGTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAATAAAGAAATGTTATTACAAGTGTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA 873  
Db 841 CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA 873

## RESULT 15

US-10-850-807-13  
; Sequence 13, Application US/10850807  
; Publication No. US20040259140A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/850,807  
; CURRENT FILING DATE: 2004-05-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni. O:36  
US-10-850-807-13

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.2e-136;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTCCTGGAATGCAACCAAGTTTAAAGAAATTCGATTATTC AAG 60  
Db 1 ATGAAAAAGTTATTATTCCTGGAATGCAACCAAGTTTAAAGAAATTCGATTATTC AAG 60  
QY 61 CTACCAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTGCAAGATAAATCTAT 120  
Db 61 CTACCAATGATTTTGTATGTTATTTAGATGTAATCAATTTTATTTGCAAGATAAATCTAT 120  
QY 121 CTTGTAATAAATGCAAAACAGTGTTTTACCCCTTAATTTCTTTTGGAGCAATCTAC 180  
Db 121 CTTGTAATAAATGCAAAACAGTGTTTTACCCCTTAATTTCTTTTGGAGCAATCTAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACCTAATTTATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACCTAATTTATGTTCTTAAT 240  
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCT 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCT 300  
QY 301 GATGCTCAATTTGGGATGATGTTTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCAATTTGGGATGATGTTTTTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360  
QY 361 TTTACGAAATTTTATTCATCAAGAAATCTACCTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTACGAAATTTTATTCATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540  
QY 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
QY 601 GAAAAAATCTTACAAAAATAAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 660  
Db 601 GAAAAAATCTTACAAAAATAAAACTATATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 660  
QY 661 ATGAACCTAGCGCCCAATTTTAAATTTCAATTTTATCATCAAGAAAAATAACTACACT 720  
Db 661 ATGAACCTAGCGCCCAATTTTAAATTTCAATTTTATCATCAAGAAAAATAACTACACT 720  
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTTATGAAAAATTTTCAAAAAATATTAAATTT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTTATGAAAAATTTTCAAAAAATATTAAATTT 780  
QY 781 AAAAAATAAATAAATAAAGAAATGTTTATTACAAAGTTTGAATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAATAAATAAATAAAGAAATGTTTATTACAAAGTTTGAATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA 873  
Db 841 CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA 873

Search completed: July 19, 2006, 15:59:47  
Job time : 998.192 secs

GenCore version 5.1.9  
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:26:04 ; Search time 141.172 Seconds  
(without alignments)  
8883.510 Million cell updates/sec

Title: US-10-734-719-13  
Perfect score: 873  
Sequence: 1 atgaaaaagtattattgc.....agcattatttcaagaaaaa 873

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 899801 segs, 718270062 residues

Total number of hits satisfying chosen parameters: 1799602

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US09\_NEW\_PUB\_seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US06\_NEW\_PUB\_seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US07\_NEW\_PUB\_seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US08\_NEW\_PUB\_seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/PCT\_NEW\_PUB\_seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US10\_NEW\_PUB\_seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US11\_NEW\_PUB\_seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US11\_NEW\_PUB\_seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/1/pubna/US60\_NEW\_PUB\_seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	70.2	8.0	13286	6	US-10-517-441-746 Sequence 746, App
C 2	69	7.9	19634	6	US-10-517-441-688 Sequence 688, App
C 3	68.8	7.9	5286	6	US-10-517-441-294 Sequence 294, App
C 4	68.8	7.9	5286	6	US-10-517-441-568 Sequence 568, App
C 5	67.2	7.7	8900	6	US-10-517-441-428 Sequence 428, App
C 6	67.2	7.7	8900	6	US-10-517-441-702 Sequence 702, App
C 7	67	7.7	13286	6	US-10-517-441-472 Sequence 472, App
C 8	66.2	7.6	5286	6	US-10-517-441-294 Sequence 294, App
C 9	66.2	7.6	5286	6	US-10-517-441-568 Sequence 568, App
C 10	64.2	7.4	19634	6	US-10-517-441-414 Sequence 414, App
C 11	62.4	7.1	2501	6	US-10-517-441-562 Sequence 562, App
C 12	62.2	7.1	10224	6	US-10-517-441-638 Sequence 638, App
C 13	62.2	7.1	11001	6	US-10-517-441-779 Sequence 779, App
C 14	61.6	7.1	8093	6	US-10-517-441-308 Sequence 308, App
C 15	61	7.0	9859	6	US-10-517-441-455 Sequence 455, App
C 16	61	7.0	9859	6	US-10-517-441-729 Sequence 729, App
C 17	60.6	6.9	2522	7	US-11-218-305-24510 Sequence 24510, A
C 18	60	6.9	8093	6	US-10-517-441-582 Sequence 582, App
C 19	60	6.9	8169	6	US-10-517-441-269 Sequence 269, App
C 20	59.8	6.8	5286	6	US-10-517-441-293 Sequence 293, App
C 21	59.8	6.8	5286	6	US-10-517-441-567 Sequence 567, App
C 22	59.4	6.8	6001	6	US-10-517-441-771 Sequence 771, App
C 23	59.2	6.8	2501	6	US-10-517-441-288 Sequence 288, App
C 24	59	6.8	10224	6	US-10-517-441-364 Sequence 364, App

C 25	59	6.8	11001	6	US-10-517-441-505 Sequence 505, App
C 26	59	6.8	16579	6	US-10-517-441-563 Sequence 563, App
C 27	58.8	6.7	12610	6	US-10-517-441-442 Sequence 442, App
C 28	58.4	6.7	5493	6	US-10-517-441-691 Sequence 691, App
C 29	58.4	6.7	8169	6	US-10-517-441-543 Sequence 543, App
C 30	58.4	6.7	12610	6	US-10-517-441-543 Sequence 543, App
C 31	58.2	6.7	5286	6	US-10-517-441-441 Sequence 441, App
C 32	58.2	6.7	10885	6	US-10-517-441-241 Sequence 241, App
C 33	58	6.6	4022	6	US-10-517-441-732 Sequence 732, App
C 34	58	6.6	8172	6	US-10-517-441-443 Sequence 443, App
C 35	58	6.6	8172	6	US-10-517-441-717 Sequence 717, App
C 36	57.8	6.6	1000	8	US-11-266-748A-281383 Sequence 281383, App
C 37	57.8	6.6	1000	8	US-11-266-748A-308023 Sequence 308023, App
C 38	57.6	6.6	7467	6	US-10-517-441-448 Sequence 448, App
C 39	57.6	6.6	7467	6	US-10-517-441-722 Sequence 722, App
C 40	57.6	6.6	8093	6	US-10-517-441-307 Sequence 307, App
C 41	57.6	6.6	8093	6	US-10-517-441-581 Sequence 581, App
C 42	57.4	6.6	4022	6	US-10-517-441-458 Sequence 458, App
C 43	57.4	6.6	8020	6	US-10-517-441-634 Sequence 634, App
C 44	57.4	6.6	12610	6	US-10-517-441-715 Sequence 715, App
C 45	57.2	6.6	9859	6	US-10-517-441-456 Sequence 456, App

## ALIGNMENTS

RESULT 1  
US-10-517-441-746/c  
; Sequence 746, Application US/10517441  
; Publication No. US20060121467A1  
; GENERAL INFORMATION:  
; APPLICANT: FOEKENS, John  
; APPLICANT: HARBECK, Nadia  
; APPLICANT: KOENIG, Thomas  
; APPLICANT: MAIER, Sabine  
; APPLICANT: MARTENS, John  
; APPLICANT: MODEL, Fabian  
; APPLICANT: NIMMICH, Inko  
; APPLICANT: RUJAN, Tamas  
; APPLICANT: SCHMITT, Armin  
; APPLICANT: SCHMITT, Manfred  
; APPLICANT: LOOK, Maxime P.  
; APPLICANT: MARX, Almuth  
; APPLICANT: HOEFLE, Heinz  
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cancer  
; FILE REFERENCE: 47675-93  
; CURRENT APPLICATION NUMBER: US/10517441  
; PRIOR FILING DATE: 2004-12-11  
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
; PRIOR FILING DATE: 2003-10-01  
; PRIOR APPLICATION NUMBER: DE 10317955.0  
; PRIOR FILING DATE: 2003-04-17  
; PRIOR APPLICATION NUMBER: DE 10300096.8  
; PRIOR FILING DATE: 2003-01-07  
; PRIOR APPLICATION NUMBER: DE 10245779.4  
; NUMBER OF SEQ ID NOS: 2147  
; SEQ ID NO 746  
; LENGTH: 13286  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-746

Query Match 8.0%; Score 70.2; DB 6; Length 13286;  
Best Local Similarity 47.1%; Pred. No. 0.0012;  
Matches 317; Conservative 0; Mismatches 348; Indels 8; Gaps 3;  
QY 175 TACTACACTTTAAACATTTTAAATCCAAATATGAGACCGCAATATATGTT 234  
DB 10482 TCCTTTATTACTAAAAAATCCAACTACCTCTTAACCCCTACATATTATCTT 10423

```
QY 235 TCTAATACACCAAGCTCACTAGAAAATGAAAAATTTTGTGTAAGAACTTTTACGATTAT 294
D 10422 ATTAAAAAACAATATTAATCAATAAATATCAATTTTATATAAAAAAATTTTAAATAAAC 10363
QY 295 TTCTCTGATGCTCATTTGGGATA--TGATTTTTTAAACAACCTTAAAGAAATTTAATGCTT 352
D 10362 TTCTATTTAAAAAATAAAATTAATTTTAACTTCTACTATTTTAAAAACAATAATTATA 10303
QY 353 ATTTTAAATTTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGTG 412
D 10302 TTTTAAATATAAACAACAATTTTAAACAACAACCTTTACCTTATTTTATATATTTCTA 10243
QY 413 CAGTAGCAGTCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAA 472
D 10242 ATATTTCATATAACATTTCCATTTAAAAAATTTTACTTTTACTCTAAATTTTCATAACTCAAT 10183
QY 473 ATGGGTCACTTATGCT--TTTGTATACCAACAAGAAATCTTTTAAACATAGCCCCCTCA 530
D 10182 ATCTAATAAAAAAACAATATTTTAAATTTTAAAAAATAAACATATAAAAAAACAATTT 10123
QY 531 TTTTAAAAATGATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTT 590
D 10122 TCTAATAATATATAAACAATAATAACAATCTATAAATAAACAACAATTAATTAACATCTTT 10063
QY 591 A-----GAATTTCTAGAAAAAATCTTACAAAAAATAAATAATTTGCTTATGCTCTCAATAGTC 646
D 10062 AACTCCTATTTTCTAAAAAATAATATTTTAAATATAAATTTCTAAAAAATAAATAAAT 10003
QY 647 TTTTAGCAATTTTATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAA 706
D 10002 AAAATCTATTTCATCAACAACAATTTAAATATACACCAATAAATTTACATTTTACATTTCAA 9943
QY 707 AAAATAACTACATAAAGATATATCTATACCTTTAGTGAGGCTTATGGAATTTTCAA 766
D 9942 ACATTTTAAATAAATTTCTAAACATTTCTTATTTTAAACCATCAATATCTTATTA 9883
QY 767 AAAATATTAATTTTAAAAAATAAATAAATAAAGAAATTTTATTTACAAAGTTGATATAAG 826
D 9882 ACACCTAAATCTTACTCAAAACATATAAATAATTTCTAAATCAATTAACATATTTTAAACAC 9823
QY 827 ATCTATTAAAGATT 839
D 9822 ATTTCAATAATAT 9810

RESULT 2
US-10-517-441-688/c
; Sequence 688, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKEN, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMWICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
```

```
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 688
; LENGTH: 19634
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-688
```

```
Query Match 7.9%; Score 69; DB 6; Length 19634;
Best Local Similarity 46.5%; Pred. No. 0.0021;
Matches 328; Conservative 0; Mismatches 370; Indels 7; Gaps 3;

QY 157 AATTTCTCTTTTGGAGCAATACACACTTTTAAACATTTTATCCAAATCAAGATATGAG 216
D 7372 AATTTCTTTATAACAAAAACAACACATAAATAAATATACTTTAAATTTTTTATATAAA 7313
QY 217 ACCGAACATTAATTTGTTCTTAATTACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGT 276
D 7312 ATATTAACCACTTAATAATACAAATTAACAAAACTTAATTAATAATTTTCTATTTATA 7253
QY 277 AAAACCTTTTACGATTAATTTTCTGATGCTCATTTGGGATATGATTTTTTAAACAACTT 336
D 7252 AAAACCAATATATTTCAATCACTTCACTATAAACAATATAAATTTCAATTTTCATAAAA 7193
QY 337 AAAGAATTTAATGCTTTTAAATTTTCAAGAAATTTTCAATCAAGAAATTTACCTCA 396
D 7192 AAAAATTTTACTTAAACACATTTTCAATTTTCTTTCCCATCAATATATATCACTAAT 7133
QY 397 GGGGCTATATGTCAGTAGCAGTAGCCATAGGATACAAAGAAATTTTATCTTTTCGGGA 456
D 7132 TTATACCAAAAAATTTTACAAAAAACAATTTCTTTTAAATTTTATATTTTATAA 7077
QY 457 ATTGATTTTATCAAAATGGGTCACTTTATGCTTTTGTATACCAACAAAGAAATCTTTTA 516
D 7076 ACTCTATATTTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7019
QY 517 AAACATAGCCCTGATTTTAAATATGATCGCTACACTATATC--GGACATAGTAAAAATAC 575
D 7018 CAATTTAACTTTTAAAAAATAACAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6959
QY 576 AGATATAAAGCTTTTAGAATTTCTAGAAAAAATTTTACAAAAAATAAATAAATAAATAA 635
D 6958 TCATTAACACCCCATAAATCTTTCTTCTATATCAATTTTAAATAAACAATAAATAAATAA 6899
QY 636 TCCTAATAGTCTTTTGGCAAAATTTTATAGAACTTAGCGCCAAATTTTAAATTTCAAAATTT 695
D 6898 ACACAAAAATACTATAATAAACCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 6839
QY 696 CATACAGAAAAAATACTACATCAATAGATATATCACTACCTTTCTAGGAGGCTTATGG 755
D 6838 ACTTAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6779
QY 756 AAAATTTTCAAAAAATATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATA 815
D 6778 CTTTATTAATAATTTTACCTTATAAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 6719
QY 816 GTTGATAAAGATCTATTAAAGATTTACCTAGTATATAAAGCAATTA 860
D 6718 AAACCTAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6674
```

```
RESULT 3
US-10-517-441-294/c
; Sequence 294, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKEN, John
; APPLICANT: HARBECK, Nadia
```

```

? APPLICANT: KOENIG, Thomas
? APPLICANT: MAIER, Sabine
? APPLICANT: MARTENS, John
? APPLICANT: MODEL, Fabian
? APPLICANT: NIMMICH, Inko
? APPLICANT: RUJAN, Tamas
? APPLICANT: SCHMITT, Armin
? APPLICANT: SCHMITT, Manfred
? APPLICANT: LOOK, Maxime P.
? APPLICANT: MARK, Almuth
? APPLICANT: HOEFLER, Heinz
? TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
? TITLE OF INVENTION: proliferative disorders
? FILE REFERENCE: 47675-93
? CURRENT APPLICATION NUMBER: US/10/517,441
? CURRENT FILING DATE: 2004-12-11
? PRIOR APPLICATION NUMBER: PCT/EP2003/010881
? PRIOR FILING DATE: 2003-10-01
? PRIOR APPLICATION NUMBER: DE 10317955.0
? PRIOR FILING DATE: 2003-04-17
? PRIOR APPLICATION NUMBER: DE 10300096.8
? PRIOR FILING DATE: 2003-01-07
? PRIOR APPLICATION NUMBER: DE 10245779.4
? PRIOR FILING DATE: 2002-10-01
? NUMBER OF SEQ ID NOS: 2147
? SEQ ID NO 294
? LENGTH: 5286
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
? US-10-517-441-294

```

Query Match	7.9%;	Score 68.8;	DB 6;	Length 5286;
Best Local Similarity	44.4%;	Pred. No. 0.0019;		
Matches	366;	Conservative	0;	Mismatches 452; Indels 6; Gaps 2;
QY	40	AAAGAAATGATTATTCTCRAAGGCTACCAATGATTTTGATGTATTAGATGTGAATCAATTT	99	
Db	5175	ATATAAATAATATTATATAATATATATATATATATATATATATATATAAATAATATAATATATA	5116	
QY	100	TATTTTGAAGATAAATACTATCTTGTAATAAAAAATGCAAAAACAGTGTTTTACACCCTAAT	159	
Db	5115	TATAATAATAATAATATATATATATATATATATATAATAATAATAATAATAATAATAATAAT	5056	
QY	160	TTCTTCTTTGGAGCAATACACACTTTAAACAATTTAAATCCAAATCAAGAATATGAGACC	219	
Db	5055	ATAATAATA	4996	
QY	220	GAACTAAATATGTGTTCTTAATTTACACACCAAGCTCATCTAGAAATGAANAATTTTGTAAA	279	
Db	4995	TATTATATATATATAATAATATATATATATATATATAATAATAATAATAATAATAATAATAAT	4936	
QY	280	ACTTTTTCAGTATATTTTCTCGATGCTCATTTGGGATATGATTTTTTTTAAACAACCTAAA	339	
Db	4935	ATATATTATATATATATAATAAT	4876	
QY	340	GAATTTTAATGCCTATTTTAAATTTTCGAAAATTTATTTCAATCAAAGAAATTTACCTCAGGG	399	
Db	4875	AATATTATAAT	4816	
QY	400	GTCTATATGTGTCAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTT	459	
Db	4815	AACAAATATAAATATATATATATATATAAATATAATATATATATAATAATAATAAT-----ATAT	4761	
QY	460	GATTTTTTATCAAAATGGGTCACTCTTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAA	519	
Db	4760	ATTATATAAATTTATATAATAATATATATATATATATATATATATATATAATAATAATAAT	4701	
QY	520	CTAGCCCTGATTTTAAAAAUGATCGCTCACACTATATCGGACATAGTAAAAATPACAGAT	579	
Db	4700	ATATTATATAAATTTTATATAATAATATATATATATATATAATAATATATATAATAATAATAATA	4641	

Qy 580 ATAAAGCTTTAGAATTTCTAGAAAAAACCATTACAAAATAAAACTATATTGCTTATGTCCT 639  
 Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 4640 TATATTAATATTATATAAATATATATTATATATAAATTTATATAAATATATATATATATT 4581  
 Qy 640 AATAGTCTTTAGCAAAATTTTAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATA 699  
 Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 4580 TAATATTATATAATATATATTATATATATTTAAATATTATA-TAATATATATTTATATATATA 4522  
 Qy 700 CAAGAAAAAATAACTACACTAAAGATATATACTCATACCTTCTAGTGAGGCTTATGGAAAA 759  
 Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 4521 TAATATTATTA 4462  
 Qy 760 TTTTCAAAAAATATTAATTTTTAAAAAATAAAATTTAAAGAAAAATGTTTATTACAAGTTG 819  
 Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 4461 TATAATATATAATA 4402  
 Qy 820 ATAAGAGTCTTAAGATTTACCTAGTGTATATAAAGCATTTATT 863  
 Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 4401 ATATATAATT 4358

RESULT 4  
 US-10-517-441-568/c  
 ; Sequence 568, Application US/10517441  
 ; Publication No. US20060121467A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FOSKENS, John  
 ; APPLICANT: HARBECK, Nadia  
 ; APPLICANT: KOENIG, Thomas  
 ; APPLICANT: MAIER, Sabine  
 ; APPLICANT: MARTENS, Joun  
 ; APPLICANT: MODEL, Fabian  
 ; APPLICANT: NIMMRICH, Inko  
 ; APPLICANT: RUJAN, Tamas  
 ; APPLICANT: SCHMITT, Armin  
 ; APPLICANT: SCHMITT, Manfred  
 ; APPLICANT: LOOK, Maxime P.  
 ; APPLICANT: MARX, Almuth  
 ; APPLICANT: HOEFLEER, Heinz  
 ; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of br  
 ; TITLE OF INVENTION: proliferative disorders  
 ; FILE REFERENCE: 47675-93  
 ; CURRENT APPLICATION NUMBER: US/10/517,441  
 ; CURRENT FILING DATE: 2004-12-11  
 ; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: DE 10317955.0  
 ; PRIOR FILING DATE: 2003-04-17  
 ; PRIOR APPLICATION NUMBER: DE 10300096.8  
 ; PRIOR FILING DATE: 2003-01-07  
 ; PRIOR APPLICATION NUMBER: DE 10245779.4  
 ; PRIOR FILING DATE: 2002-10-01  
 ; NUMBER OF SEQ ID NOS: 2147  
 ; SEQ ID NO 568  
 ; LENGTH: 5286  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-517-441-568

Query Match	7.9%	Score 68.8;	DB 6;	Length 5286;
Best Local Similarity	44.4%;	Pred. No. 0.0019;		
Matches 366;	Conservative 0;	Mismatches 452;	Indels 6;	Gaps 2;
Y	40	AAAGAAATTCATTATTCACGGCTACCAAATGATTTTGATGCTATTAGATGTAATCAATTT	99	
b	5175	ATATAAAATATATTATAATAATATATTATATTATATATTATAATATAATATATATTA	5116	
Y	100	TATTTTGAAGATAAATACTATCTTGGTAAAAAATGCAAAACAGTGTGTTTACACCCCTAAT	159	
b	5115	TATAATATAATATAATATTTATTTATTTATTTATTTATATAATATAATAATATTTATATAATAT	5056	







[illegible]



APPLICANT: MARTENS, John  
APPLICANT: MODEL, Fabian  
APPLICANT: NIMMICH, Inko  
APPLICANT: RUJAN, Tamas  
APPLICANT: SCHMITT, Armin  
APPLICANT: SCHMITT, Manfred  
APPLICANT: LOOK, Maxime P.  
APPLICANT: MARX, Almuth  
APPLICANT: HOEFER, Heinz  
TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
TITLE OF INVENTION: proliferative disorders  
FILE REFERENCE: 47675-93  
CURRENT APPLICATION NUMBER: US/10/517,441  
CURRENT FILING DATE: 2004-12-11  
PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
PRIOR FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: DE 10317955.0  
PRIOR FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: DE 10300096.8  
PRIOR FILING DATE: 2003-01-07  
PRIOR APPLICATION NUMBER: DE 10245779.4  
PRIOR FILING DATE: 2002-10-01  
NUMBER OF SEQ ID NOS: 2147  
SEQ ID NO 562  
LENGTH: 2501  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-562

Query Match 7.1%; Score 62.4; DB 6; Length 2501;  
Best Local Similarity 49.5%; Pred. No. 0.023;  
Matches 190; Conservative 0; Mismatches 191; Indels 3; Gaps 1;  
QY 492 TGATACCAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAATGATCGCTCA 551  
Db 1744 TAATAAATAACATATAATAATAATAATAATAATAATAATAATAATAATA 1685  
QY 552 CTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTAGAAAAA 611  
Db 1684 CTTAAACAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATA 1625  
QY 612 CAAATAAACTATATGCTTATGCTTAATAGTCTTTTAGCAATTTTATAGAACTAGC 671  
Db 1624 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1565  
QY 672 GCCAAATTTAAATTTTATCATTA--CAAGAAAAAATAACTACACTAAAGATAT 728  
Db 1564 AATAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1505  
QY 729 ACTCATACCTTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTTAAAAA 788  
Db 1504 AATATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1445  
QY 789 AAAAAATTAAGAAATGTTTATTACAGCTTATGATAAAGATCTATTAAAGATTACCTAGTGA 848  
Db 1444 AACAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1385  
QY 849 TATAAGCATTTTCAAGGAAA 872  
Db 1384 AATAAATAAACAATAAAAAATA 1361

RESULT 12  
US-10-517-441-638/c  
Sequence 638, Application US/10517441  
Publication No. US20060121467A1  
GENERAL INFORMATION:  
APPLICANT: FOEKEN, John  
APPLICANT: HARBECK, Nadia  
APPLICANT: KOENIG, Thomas  
APPLICANT: MAIER, Sabine

APPLICANT: MARTENS, John  
APPLICANT: MODEL, Fabian  
APPLICANT: NIMMICH, Inko  
APPLICANT: RUJAN, Tamas  
APPLICANT: SCHMITT, Armin  
APPLICANT: SCHMITT, Manfred  
APPLICANT: LOOK, Maxime P.  
APPLICANT: MARX, Almuth  
APPLICANT: HOEFER, Heinz  
TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
TITLE OF INVENTION: proliferative disorders  
FILE REFERENCE: 47675-93  
CURRENT APPLICATION NUMBER: US/10/517,441  
CURRENT FILING DATE: 2004-12-11  
PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
PRIOR FILING DATE: 2003-10-01  
PRIOR APPLICATION NUMBER: DE 10317955.0  
PRIOR FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: DE 10300096.8  
PRIOR FILING DATE: 2003-01-07  
PRIOR APPLICATION NUMBER: DE 10245779.4  
PRIOR FILING DATE: 2002-10-01  
NUMBER OF SEQ ID NOS: 2147  
SEQ ID NO 638  
LENGTH: 10224  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-517-441-638

Query Match 7.1%; Score 62.2; DB 6; Length 10224;  
Best Local Similarity 45.0%; Pred. No. 0.03;  
Matches 311; Conservative 0; Mismatches 378; Indels 2; Gaps 2;  
QY 182 CTTTAAACATTTAATCCAAATCAAGAATATGAGACCGAACTAATATGTGTCTTAAT 241  
Db 7181 CTCAAAAAATAAATAAATCAACATAAATAAATAAATAAATAAATAAATAAATA 7122  
QY 242 ACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCCCTG 301  
Db 7121 AAAAAAACAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7062  
QY 302 ATGCTCATTTGGGATATGATTTTAAAACAACTTAAAGAAATTAATGCTTATTTAAAT 361  
Db 7061 AACCTAATAACCATTTTACTACAAAAAATAAATAAATAAATAAATAAATA 7002  
QY 362 TTCAGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGTGAGTAGCCA 421  
Db 7001 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6942  
QY 422 TAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 481  
Db 6941 AATCTACACAAAAAATACTAATCTAAAAATTTTAAATCAATATTTTCAACCATTAAT 6882  
QY 482 CTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAATG 541  
Db 6881 AACAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6822  
QY 542 ATCGCTCACATATATCGGACATAGTAAAAATACAGATAATAAAGCTTTTGAATTTCTAG 601  
Db 6821 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6763  
QY 602 AAAAACTTACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 661  
Db 6762 AAAAACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6704  
QY 662 TAGAATAGCGCGCAATTTTAAATTTTAAATTTTATCATCAAGAAAAAATAAATAAATA 721  
Db 6703 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6644  
QY 722 AAGATATATCTCATACCTTCTAGTAGGCTTATGGAATTTTCAAAAAATTTTAAATTTT 781



Matches 283; Conservative 0; Mismatches 324; Indels 7; Gaps 2;

QY 209 AATATGAGACCGAAGTAATATGTTCTTAATACAAACGAGCTCATCTAGAAATGAAA 268  
 |||||  
 Db 2420 AATAACCCCTCTTAATATTTTCTATCTTTTCGAAACGACCTTCGAAATACCA 2361  
 |||||  
 QY 269 ATTTTGTAAACATTTTACAGATTTTCTGATGCTCATTTGGGATATGATTTTTTA 328  
 |||||  
 Db 2360 AATTTTATAATCATCTCTTAAATTTTAAACCTTATAACTATAACAAAAATATAAAAAA 2301  
 |||||  
 QY 329 AACAACTTTAAAGATTTAATGCTTATTTTAAATTTTCAAGAAATTTTCAATCAAGAA 388  
 |||||  
 Db 2300 AAAATTTTAAACATATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2241  
 |||||  
 QY 389 TTACCTCAGGGTCTATATGTTGTCAGTAGCCCTAGGATACAAAGAAATTTATC 448  
 |||||  
 Db 2240 ACCCAAAATACACCTTAAGCTTATATCAATTTATTTTAAAT-CAATTATTAATAAC 2182  
 |||||  
 QY 449 TTTCGGGAATGATTTTATCAAAATGGGTGATCTTATGCTTTTGTATACCAAGAA 508  
 |||||  
 Db 2181 GTCCGTAGCTTAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2122  
 |||||  
 QY 509 ATCTTTTAAACTAGCCCTGATTTTAAATATGATCGCTCACATATATCGGACATCTA 568  
 |||||  
 Db 2121 TTATTTAAATATTAATACTTCAAAACAATAAATAAATAAATAAATAAATAA 2068  
 |||||  
 QY 569 AAAATACAGATATAAAGCTTTTAGAATTTCTAGAAAAAACTTACAAAAATAAATACTATAT 628  
 |||||  
 Db 2067 AAAAAACCCTAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2008  
 |||||  
 QY 629 GCTTATGCTTAATAGTCTTTTAGCAATTTTATAGAACTAGCGCAAAATTTAAATCAA 688  
 |||||  
 Db 2007 CTAAAAATATTTTCTCTTTTCACTTATATAATAATAAATAAATAAATAAATAA 1948  
 |||||  
 QY 689 ATTTTATACACAGAAAAAATAAATACTACATTAAGATATATCTATACCTTCTAGTGAGG 748  
 |||||  
 Db 1947 TTTTATTTCTACTATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1888  
 |||||  
 QY 749 CTATGGAATAATTTTCAAAAAATATTAATTTTAAAAAATAAATAAATAAATAAATAA 808  
 |||||  
 Db 1887 CCAAAATACATATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1828  
 |||||  
 QY 809 ATTACAAGTTGATA 822  
 |||||  
 Db 1827 AATAAAAAATTTTA 1814

RESULT 15

US-10-517-441-455/c  
 ; Sequence 455, Application US/10517441  
 ; Publication No. US20060121467A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FOSKENS, John  
 ; APPLICANT: HARBECK, Nadia  
 ; APPLICANT: KOENIG, Thomas  
 ; APPLICANT: MAIER, Sabine  
 ; APPLICANT: MARTENS, John  
 ; APPLICANT: MODEL, Fabian  
 ; APPLICANT: NIMMICH, Inko  
 ; APPLICANT: RUJAN, Tamas  
 ; APPLICANT: SCHMITT, Armin  
 ; APPLICANT: SCHMITT, Manfred  
 ; APPLICANT: LOOK, Maxime P.  
 ; APPLICANT: MARX, Almut  
 ; APPLICANT: HOEFER, Heinz  
 ; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel  
 ; TITLE OF INVENTION: proliferative disorders  
 ; FILE REFERENCE: 47675-93  
 ; CURRENT APPLICATION NUMBER: US/10/517,441  
 ; CURRENT FILING DATE: 2004-12-11  
 ; PRIOR APPLICATION NUMBER: PCT/BP2003/010891  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: DE 10317955.0

; PRIOR FILING DATE: 2003-04-17  
 ; PRIOR APPLICATION NUMBER: DE 10300096.8  
 ; PRIOR FILING DATE: 2003-01-07  
 ; PRIOR APPLICATION NUMBER: DE 10245779.4  
 ; PRIOR FILING DATE: 2002-10-01  
 ; NUMBER OF SEQ ID NOS: 2147  
 ; SEQ ID NO 455  
 ; LENGTH: 9859  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-517-441-455

Query Match 7.0%; Score 61; DB 6; Length 9859;  
 Best Local Similarity 50.2%; Pred. No. 0.048;  
 Matches 151; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 494 ATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAATGATCGCTCACACT 553  
 |||||  
 Db 9102 ATACCAATTTAAAAAATACATAAATATCCATTTAAAAATAATTTAAAAAATCTCTAAAAT 9043  
 |||||  
 QY 554 ATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAATTACA 613  
 |||||  
 Db 9042 AACTATAATAATTTACATATATCTATAATACACTAAAAAACCCCTAAATTTACCCCT 8983  
 |||||  
 QY 614 AAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAATTTTATAGAACTAGCGC 673  
 |||||  
 Db 8982 TAAATTAACAACCTTATATACTATATAAATTTATCTCAATTAACCTACTTTTAAAAA 8923  
 |||||  
 QY 674 CAAATTTAAATTTCAATTTTATCATAACAAGAAAAAATAAATACTACACTTAAAGATATAC 733  
 |||||  
 Db 8922 TAAAAAATCTCTAATTTTCTTAAAAAATAAACCCTACAATAACAATATATA 8863  
 |||||  
 QY 734 TACCTTCTAGTGGGCTTATGAAAAATTTTCAAAAAATTTTAAAAAATAAATAA 793  
 |||||  
 Db 8862 AATAATCCAAAAATAACCTATAAATTTAAAAATAAATAAATAAATAAATAAATAA 8803  
 |||||  
 QY 794 T 794  
 Db 8802 T 8802

Search completed: July 19, 2006, 15:44:55  
 Job time : 141.172 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:30:08 ; Search time 107.972 Seconds  
(without alignments)  
1232.267 Million cell updates/sec

Title: US-10-734-719-14  
Perfect score: 1564  
Sequence: 1 MKVLIAGNPSLKBIDYSR.....KLIKDLRLPSLKIHYFKGK 291

Scoring table: BLOSUM62

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_8.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*  
10: Geneseqp2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1564	100.0	291	6 ABJ18485	Abj18485 Campyloba
2	1559	99.7	291	3 AAY97212	Aay97212 Campyloba
3	1559	99.7	291	6 ABJ18482	Abj18482 Campyloba
4	1553	99.3	291	6 ABJ18484	Abj18484 Campyloba
5	1524	97.4	291	3 AAY97210	Aay97210 Campyloba
6	1524	97.4	291	6 ABJ18480	Abj18480 Campyloba
7	1521	97.3	291	3 AAY97204	Aay97204 Campyloba
8	1521	97.3	291	6 ABJ18479	Abj18479 Campyloba
9	1510	96.5	291	3 AAY97211	Aay97211 Campyloba
10	1510	96.5	291	6 ABJ18481	Abj18481 Campyloba
11	804	51.4	294	3 AAY97215	Aay97215 Campyloba
12	804	51.4	294	6 ABJ18483	Abj18483 Campyloba
13	739	47.3	430	2 AAY45221	Aay45221 Campyloba
14	688.5	44.0	303	6 ABR40195	Abri40195 alpha-2,3
15	119.5	7.6	958	8 ADN19559	Adn19559 Bacterial
16	115	7.4	2133	8 ADP25432	Adp25432 Plasmodiu
17	114	7.3	392	3 AAB15991	Aab15991 E. coli p
18	111.5	7.1	908	6 ABU19261	Abu19261 Protein e
19	111	7.1	1817	3 AAB18301	Aab18301 Plasmodiu
20	110	7.0	863	6 ADB08930	Adb08930 Allostoc
21	109.5	7.0	1121	3 AAB18241	Aab18241 Plasmodiu
22	109	7.0	1398	3 AAB18292	Aab18292 Plasmodiu
23	108.5	6.9	1062	8 ADP25429	Adp25429 Plasmodiu

## ALIGNMENTS

### RESULT 1

ABJ18485  
ID ABJ18485 standard; protein; 291 AA.  
XX  
AC ABJ18485;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni bifunctional sialtransferase cstII #6.  
XX  
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW Cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.  
XX  
PN WO200274942-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 22-FEB-2002; 2002WO-CA000229.  
XX  
PR 21-MAR-2001; 2001US-00816028.  
XX  
(CANA ) NAT RES COUNCIL CANADA.  
PI Gilbert M, Wakarchuk WW;  
XX  
XX WPI; 2003-040554/03.  
DR N-PSDB; ABT13671.  
XX

24 108.5 6.9 1284 6 ABU48910  
25 108 6.9 822 5 ABB77605  
26 108 6.9 1021 2 AAW98555  
27 108 6.9 1802 3 AAB18217  
28 107.5 6.9 295 7 ADC96625  
29 107.5 6.9 400 6 ABU25530  
30 107 6.8 1815 8 ADP25442  
31 106.5 6.8 305 6 ABM69267  
32 106.5 6.8 1294 7 ADC21264  
33 106 6.8 1182 3 AAB18288  
34 106 6.8 2539 3 AAB18198  
35 106 6.8 3973 3 AAB18253  
36 105.5 6.7 615 6 ABU25054  
37 105.5 6.7 772 4 ABB57750  
38 105.5 6.7 1346 3 AAB18236  
39 105.5 6.7 1465 6 ABU19347  
40 105 6.7 1817 3 AAB18255  
41 105 6.7 2013 3 AAB18265  
42 104.5 6.7 388 8 ADK16435  
43 104.5 6.7 422 4 AAU35736  
44 104.5 6.7 422 6 ABR83248  
45 104.5 6.7 422 6 ABR83247

Abu48910 Protein e  
Abb77605 Anepv sec  
Aaw98555 H. pylori  
Aab18217 Plasmodiu  
Adc96625 E. faeciu  
Abu25530 Protein e  
Adp25442 Plasmodiu  
Abm69267 Photorhab  
Adc21264 Plasmodiu  
Aab18288 Plasmodiu  
Aab18198 Plasmodiu  
Aab18253 Plasmodiu  
Abu25054 Protein e  
Abb57750 Drosophil  
Aab18236 Plasmodiu  
Abu19347 Protein e  
Aab18255 Plasmodiu  
Aab18265 Plasmodiu  
Adk16435 Nanoarcha  
Aau35736 Helicobac  
Abr83248 H. pylori  
Abr83247 H. pylori

New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.

Disclosure; Page 99; 107pp; English.

The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an: acyltransferase; glycosyltransferase; GalNAc (N-Acetylgalactosamine) transferase; galactosyltransferase; galactosyltransferase; sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein



sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present amino acid sequence represents a Campylobacter jejuni protein of the invention

Sequence 291 AA;

Query Match 100.0%; Score 1564; DB 6; Length 291;  
Best Local Similarity 100.0%; Pred. No. 1.9e-147;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYILGKKCKTVFTNPFPEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYILGKKCKTVFTNPFPEQYY 60  
QY 61 TLKHLIQNOEYETELIMCSNMQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
DB 61 TLKHLIQNOEYETELIMCSNMQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIFYNQRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 FHEIFYNQRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHVIGHSKNTDIKALEFLEKTYKIYLCPCNSLIANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHVIGHSKNTDIKALEFLEKTYKIYLCPCNSLIANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 2

AA97212  
ID AA97212 standard; protein; 291 AA.  
XX  
AC AA97212;  
XX  
DT 12-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE Campylobacter jejuni O:19 serotype CstII sialyltransferase.  
XX  
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside.  
XX  
OS Campylobacter jejuni; O:19 serotype.  
XX  
FN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CA000086.  
XX  
PR 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
XX (CANADA) NAT RES COUNCIL CANADA.  
PA  
XX Gilbert M, Wakarchuk WM;  
XX  
XX WPI; 2000-524418/47.  
DR N-PSDB; AAA53726.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for biosynthesis of ganglioside and ganglioside mimics, as diagnostic

PT reagents and as immunogen for producing antibodies.

XX Disclosure; Page 96-97; 120pp; English.

PS A reaction mixture for the synthesis of a sialylated oligosaccharide is useful for synthesising sialylated oligosaccharides such as ganglioside, lysoganglioside or their mimics. Glycosyltransferases are useful for chemo-enzymatic synthesis of oligosaccharides, including gangliosides and other oligosaccharides that have biological activity. The enzymes and nucleic acids that encode them are useful for studies of the pathogenesis mechanisms of organisms that synthesize ganglioside mimics, such as C. jejuni and the nucleic acids are used as probes to study expression of genes involved in ganglioside mimetic synthesis. Antibodies raised against the glycosyltransferases are also useful for analyzing the expression patterns of these genes involved in pathogenesis. The nucleic acids are also useful for designing antisense oligonucleotides for inhibiting expression of the Campylobacter enzymes that are involved in the biosynthesis of ganglioside mimics that can mask the pathogens from the host's immune system. The oligosaccharides are useful as diagnosing reagents or as therapeutics and as immunogens for producing antibodies. Bacterial glycosyltransferase can be used to catalyse the formation of oligosaccharides that are identical to the corresponding mammalian structures and are easier and less expensive to produce in large quantity, compared to the mammalian glycosyltransferase. The bacterial origin of the enzymes facilitates expression of large quantities of the enzymes using relatively inexpensive prokaryotic expression systems.

CC (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 291 AA;

SQ Query Match 99.7%; Score 1559; DB 3; Length 291;  
Best Local Similarity 99.7%; Pred. No. 6e-147;  
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYILGKKCKTVFTNPFPEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYILGKKCKAVFTNPFPEQYY 60  
QY 61 TLKHLIQNOEYETELIMCSNMQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
DB 61 TLKHLIQNOEYETELIMCSNMQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIFYNQRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
DB 121 FHEIFYNQRITSGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
QY 181 DRSHVIGHSKNTDIKALEFLEKTYKIYLCPCNSLIANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHVIGHSKNTDIKALEFLEKTYKIYLCPCNSLIANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 3

ABJ18482  
ID ABJ18482 standard; protein; 291 AA.  
XX  
AC ABJ18482;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni bifunctional sialtransferase cstII #4.  
XX  
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.

XX WO200274942-A2.  
XX 26-SEP-2002.  
XX 22-FEB-2002; 2002WO-CA000229.  
XX 21-MAR-2001; 2001US-00816028.  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX Gilbert M, Wakarchuk WW;  
XX WPI; 2003-040554/03.  
XX N-PSDB; ABT13669.  
XX  
XX New glycosyltransferases from Campylobacter, useful for synthesizing  
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
XX Disclosure; Page 99; 107pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of  
XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
XX may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
XX Acetylgalactosamine) transferase; galactosyltransferase;  
XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
XX sequences of the invention are useful for ganglioside synthesis, studying  
XX ganglioside mimetics, and for designing oligonucleotides to inhibit  
XX expression of Campylobacter enzymes involved in the biosynthesis of  
XX ganglioside mimetics that can mask the pathogen's from the host's immune  
XX system. The C. jejuni oligosaccharides of the invention may be used as  
XX diagnostic reagents (e.g. to locate areas of inflammation or tumour  
XX metastasis). The present amino acid sequence represents a Campylobacter  
XX jejuni protein of the invention  
XX  
XX Sequence 291 AA;  
XX  
XX Query Match 99.7%; Score 1559; DB 6; Length 291;  
XX Best Local Similarity 99.7%; Pred. No. 6e-147;  
XX Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLLGKCKTAVYTPNFFFEQYY 60  
XX 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLLGKCKAVYTPNFFFEQYY 60  
XX  
XX 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDPDAHLGYDFFKQLKFNAYFK 120  
XX 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDPDAHLGYDFFKQLKFNAYFK 120  
XX  
XX 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180  
XX 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180  
XX  
XX 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
XX 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
XX  
XX 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
XX 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
XX  
XX RESULT 4  
XX ABJ18484  
XX ID ABJ18484 standard; protein; 291 AA.  
XX  
XX AC ABJ18484;  
XX  
XX DT 07-FEB-2003. (first entry)  
XX  
XX DE Campylobacter jejuni bifunctional sialtransferase cstII #5.

XX Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
XX GalNAC transferase; N-Acetylgalactosamine transferase;  
XX galactosyltransferase; sialyltransferase; sialic acid synthase;  
XX cytidine 5'-monophosphate sialic acid synthetase;  
XX CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
XX ganglioside mimetics; inflammation; tumour metastasis.  
XX  
XX Campylobacter jejuni.  
XX  
XX WO200274942-A2.  
XX 26-SEP-2002.  
XX  
XX 22-FEB-2002; 2002WO-CA000229.  
XX  
XX 21-MAR-2001; 2001US-00816028.  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX WPI; 2003-040554/03.  
XX N-PSDB; ABT13670.  
XX  
XX New glycosyltransferases from Campylobacter, useful for synthesizing  
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
XX Disclosure; Page 99; 107pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of  
XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
XX may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
XX Acetylgalactosamine) transferase; galactosyltransferase;  
XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
XX sequences of the invention are useful for ganglioside synthesis, studying  
XX ganglioside mimetics, and for designing oligonucleotides to inhibit  
XX expression of Campylobacter enzymes involved in the biosynthesis of  
XX ganglioside mimetics that can mask the pathogen's from the host's immune  
XX system. The C. jejuni oligosaccharides of the invention may be used as  
XX diagnostic reagents (e.g. to locate areas of inflammation or tumour  
XX metastasis). The present amino acid sequence represents a Campylobacter  
XX jejuni protein of the invention  
XX  
XX Sequence 291 AA;  
XX  
XX Query Match 99.3%; Score 1553; DB 6; Length 291;  
XX Best Local Similarity 99.3%; Pred. No. 2.4e-146;  
XX Matches 289; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLLGKCKTAVYTPNFFFEQYY 60  
XX 1 MKKVIIAGNGPSLKEIDYSRLPNDVFRNCQFYFEDKYLLGKCKAVYTPNFFFEQYY 60  
XX  
XX 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKQLKFNAYFK 120  
XX 61 TLKHLIQOEYETELIMCSNNOAHLENENFVKTFYDYPDAHLGYDFFKQLKFNAYFK 120  
XX  
XX 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180  
XX 121 FHEIYNQRIITSGVVMCAVAIALGYKEIYLSGIDFYONGSSYAFDTKQENLLKLAPDPKN 180  
XX  
XX 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
XX 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
XX  
XX 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
XX 241 KDILIPSSSEAYGKFSKNINFKKIKIKENYVYKLIKDLLRLPSDIKHYPFGK 291  
XX



XX Sequence 291 AA;  
SQ  
Query Match 97.4%; Score 1524; DB 6; Length 291;  
Best Local Similarity 97.6%; Pred. No. 1.9e-143;  
Matches 284; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSLPNDVFRNQYFEDKYLGGKCKTVYTPNPFEEQY 60  
DB 1 MKKVIAGNGPSLKEIDYSLPNDVFRNQYFEDKYLGGKCKTVYTPNPFEEQY 60  
QY 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFYDYPDAHLGYDFPKQKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFYDYPDAHLGYDFPKQKEFNAYFK 120  
QY 121 FHEIYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFN 180  
DB 121 FHEIYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKENYVYKLIKDLLRLPSDIKHYPFKG 240  
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKENYVYKLIKDLLRLPSDIKHYPFKG 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKG 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKG 291  
RESULT 7  
AAAY97204  
ID AAAY97204 standard; protein; 291 AA.  
XX  
AC AAAY97204;  
XX  
DT 12-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE Campylobacter jejuni OH4384 CstII sialyltransferase.  
XX  
DE Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside.  
XX  
OS Campylobacter jejuni; OH4384.  
XX  
FN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CA000086.  
XX  
PF 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
PA (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX  
DR WPI; 2000-524418/47.  
DR N-PSDB; AAA53721, AAA53720.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
PT reagents and as immunogen for producing antibodies.  
XX  
FS Claim 13; Page 91-92; 120pp; English.  
XX  
CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and

CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 291 AA;  
Query Match 97.3%; Score 1521; DB 3; Length 291;  
Best Local Similarity 96.9%; Pred. No. 3.8e-143;  
Matches 282; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSLPNDVFRNQYFEDKYLGGKCKTVYTPNPFEEQY 60  
DB 1 MKKVIAGNGPSLKEIDYSLPNDVFRNQYFEDKYLGGKCKTVYTPNPFEEQY 60  
QY 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFYDYPDAHLGYDFPKQKEFNAYFK 120  
DB 61 TLKHLIQOEYETELIMCSNYQAHLNENFVKTFYDYPDAHLGYDFPKQKEFNAYFK 120  
QY 121 FHEIYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFN 180  
DB 121 FHEIYNQRITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDPFN 180  
QY 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKENYVYKLIKDLLRLPSDIKHYPFKG 240  
DB 181 DRSHYIGHSKNTDIIKALEFLEKTYKIKENYVYKLIKDLLRLPSDIKHYPFKG 240  
QY 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKG 291  
DB 241 KDILIPSSAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHYPFKG 291  
RESULT 8  
ABJ18479  
ID ABJ18479 standard; protein; 291 AA.  
XX  
AC ABJ18479;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni bifunctional sialtransferase cstII #1.  
XX  
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.  
XX  
PN WO200274942-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 22-FEB-2002; 2002WO-CA000229.  
XX  
PR 21-MAR-2001; 2001US-00816028.

XX PA (CANA ) NAT RES COUNCIL CANADA.  
XX PI Gilbert M, Wakarchuk WW;  
XX CDS  
XX WPI; 2003-040554/03.  
XX N-PSDB; AET13666.  
XX  
XX New glycosyltransferases from Campylobacter, useful for synthesizing  
XX gangliosides and ganglioside mimetics, and in studying the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
XX Claim 5; Page 96; 107pp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of  
XX Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
XX may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
XX Acetyl-galactosamine) transferase; galactosyltransferase;  
XX sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
XX sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
XX sequences of the invention are useful for ganglioside synthesis, studying  
XX ganglioside mimetics, and for designing oligonucleotides to inhibit  
XX expression of Campylobacter enzymes involved in the biosynthesis of  
XX ganglioside mimetics that can mask the pathogen's from the host's immune  
XX system. The C. jejuni oligosaccharides of the invention may be used as  
XX diagnostic reagents (e.g. to locate areas of inflammation or tumour  
XX metastasis). The present amino acid sequence represents a Campylobacter  
XX jejuni protein of the invention  
XX  
XX Sequence 291 AA;  
XX  
XX Query Match 97.3%; Score 1521; DB 6; Length 291;  
XX Best Local Similarity 96.9%; Pred. No. 3.8e-143;  
XX Matches 282; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MKKVIIAGNGPSLKEIDYSRLENDFVPRCNQFYEDKYLGKKCKTVFYTPNFFEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSRLENDFVPRCNQFYEDKYLGKKCKAVFYNPILFFEQYY 60  
Qy 61 TLKHLIQNEVETELIMCSNNOAHLENENFVKTYDYFPDAHLGYDFPKQKFNAYFK 120  
Db 61 TLKHLIQNEVETELIMCSNNOAHLENENFVKTYDYFPDAHLGYDFPKQKFNAYFK 120  
Qy 121 FHEIYFNQRTSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYFNQRTSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKKNYT 240  
Db 181 DSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKKNYT 240  
Qy 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLRLPSDINKHYFKGK 291  
Db 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLRLPSDINKHYFKGK 291

RESULT 9  
AAAY97211  
ID AAAY97211 standard; protein; 291 AA.  
XX AC  
XX AAAY97211;  
XX  
XX 12-SEP-2003 (revised)  
XX 22-DEC-2000 (first entry)  
XX  
XX Campylobacter jejuni O:41 serotype alpha-2,3-sialyltransferase.  
XX  
XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
XX acetyltransferase; glycosyltransferase; Beta-1,4-GalNac transferase;  
XX Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
XX sialic acid synthetase; CMP-sialic acid synthetase; mimic; antibody;  
XX immunity; immunogen; ganglioside.  
XX

OS Campylobacter jejuni; O:41 serotype.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..876  
XX FT /\*tag= a  
XX FT /product= "alpha-2,3-sialyltransferase"  
XX  
XX WO200046379-A1.  
XX  
XX 10-AUG-2000.  
XX  
XX 01-FEB-2000; 2000WO-CA000086.  
XX  
XX 01-FEB-1999; 99US-0118213P.  
XX 31-JAN-2000; 2000US-00495406.  
XX  
XX (CANA ) NAT RES COUNCIL CANADA.  
XX  
XX Gilbert M, Wakarchuk WW;  
XX WPI; 2000-524418/47.  
XX N-PSDB; AAAS3725.  
XX  
XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
XX reagents and as immunogen for producing antibodies.  
XX  
XX Claim 13; Page 94-95; 120pp; English.  
XX  
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
XX useful for synthesizing sialylated oligosaccharide such as ganglioside,  
XX lysoganglioside or their mimics. Glycosyltransferases are useful for  
XX chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
XX other oligosaccharides that have biological activity. The enzymes and  
XX nucleic acids that encode them are useful for studies of the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimics, such as C.  
XX jejuni and the nucleic acids are used as probes to study expression of  
XX genes involved in ganglioside mimetic synthesis. Antibodies raised  
XX against the glycosyltransferases are also useful for analyzing the  
XX expression patterns of these genes involved in pathogenesis. The nucleic  
XX acids are also useful for designing antisense oligonucleotides for  
XX inhibiting expression of the Campylobacter enzymes that are involved in  
XX the biosynthesis of ganglioside mimics that can mask the pathogens from  
XX the host's immune system. The oligosaccharides are useful as diagnosing  
XX reagents or as therapeutics and as immunogens for producing antibodies.  
XX Bacterial glycosyltransferase can be used to catalyse the formation of  
XX oligosaccharides that are identical to the corresponding mammalian  
XX structures and are easier and less expensive to produce in large  
XX quantity, compared to the mammalian glycosyltransferase. The bacterial  
XX origin of the enzymes facilitates expression of large quantities of the  
XX enzymes using relatively inexpensive prokaryotic expression systems.  
XX (Updated on 12-SEP-2003 to standardise OS field)  
XX  
XX Sequence 291 AA;  
XX  
XX Query Match 96.5%; Score 1510; DB 3; Length 291;  
XX Best Local Similarity 95.5%; Pred. No. 4.7e-142;  
XX Matches 278; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MKKVIIAGNGPSLKEIDYSRLENDFVPRCNQFYEDKYLGKKCKTVFYTPNFFEQYY 60  
Db 1 MKKVIIAGNGPSLKEIDYSRLENDFVPRCNQFYEDKYLGKKCKAVFYNPILFFEQYY 60  
Qy 61 TLKHLIQNEVETELIMCSNNOAHLENENFVKTYDYFPDAHLGYDFPKQKFNAYFK 120  
Db 61 TLKHLIQNEVETELIMCSNNOAHLENENFVKTYDYFPDAHLGYDFPKQKFNAYFK 120  
Qy 121 FHEIYFNQRTSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Db 121 FHEIYFNQRTSGVYMCVAIAIYALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
Qy 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKKNYT 240

Db 181 DSHYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINPKIKIKENVYKLIKOLLRLPSDIKHYPGK 291  
 Db 241 KDILIPSSAYGKFTKNINPKIKIKENIYKLIKOLLRLPSDIKHYPGK 291

RESULT 10  
 ABJ18481  
 ID ABJ18481 standard; protein; 291 AA.  
 XX  
 AC ABJ18481;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE Campylobacter jejuni bifunctional sialtransferase cstII #3.  
 XX  
 KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 XX  
 PN WO200274942-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 22-FEB-2002; 2002WO-CA000229.  
 XX  
 PR 21-MAR-2001; 2001US-00816028.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Gilbert M, Wakarchuk WW;  
 XX  
 DR WPI; 2003-040554/03.  
 XX  
 DR N-PSDB; ABT13668.  
 XX  
 PT New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Claim 5; Page 97-98; 107pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present amino acid sequence represents a Campylobacter  
 CC jejuni protein of the invention  
 XX  
 SQ Sequence 291 AA;

Query Match 96.5%; Score 1510; DB 6; Length 291;  
 Best Local Similarity 95.5%; Pred. No. 4.7e-142;  
 Matches 278; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKKVIIAGNPSLKEIDYSRPNDFVRCNQFYEDKYYLGGKCKTVYTNPFEEQY 60  
 Db 1 MKKVIIAGNPSLKEIDYSRPNDFVRCNQFYEDKYYLGGKCKTVYTNPFEEQY 60  
 QY 61 TLKHLIQNQEYETELIMCSNQNQAHLNENFVKTFDYFPDAHLGYDFPKQLKEFNAYFK 120

Db 61 TLKHLIQNQEYETELIMCSNQNQAHLNENFVKTFDYFPDAHLGYDFPKQLKEFNAYFK 120  
 QY 121 FHEIFYNQRTISGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLKLAPDFKN 180  
 Db 121 FHEIFYNQRTISGVYMCVAIAIAGYKEIYLSGIDFYQNGSSYAFDTKQENLKLAPDFKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240  
 Db 181 DSHYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIQEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINPKIKIKENVYKLIKOLLRLPSDIKHYPGK 291  
 Db 241 KDILIPSSAYGKFTKNINPKIKIKENIYKLIKOLLRLPSDIKHYPGK 291

RESULT 11  
 AAY97215  
 ID AAY97215 standard; protein; 294 AA.  
 XX  
 AC AAY97215;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 22-DEC-2000 (first entry)  
 XX  
 DE Campylobacter jejuni NCTC11168 alpha-2,3-sialyltransferase.  
 XX  
 KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
 KW immunity; immunogen; ganglioside.  
 XX  
 OS Campylobacter jejuni; NCTC11168.  
 XX  
 PN WO200046379-A1.  
 XX  
 PD 10-AUG-2000.  
 XX  
 PF 01-FEB-2000; 2000WO-CA000086.  
 XX  
 PR 01-FEB-1999; 99US-0118213P.  
 PR 31-JAN-2000; 2000US-00495406.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Gilbert M, Wakarchuk WW;  
 XX  
 DR WPI; 2000-524418/47.  
 XX  
 PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 PT reagents and as immunogen for producing antibodies.  
 XX  
 PS Disclosure; Page 97-98; 120pp; English.  
 XX  
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesising sialylated oligosaccharides such as ganglioside,  
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyse the formation of



production of sialic acid-containing complex sugars.

Claim 9; Page 60-62; 66pp; Japanese.

The present invention relates to a method for producing alpha-2,3/alpha-2,8-stalyltransferase. alpha-2,3/alpa-2,8-sialyltransferase can be used in a method for producing sialic acid-containing complex sugars from an oligosaccharide substrate using a culture of a microorganism transformed with DNA encoding alpha-2,3/alpa-2,8-stalyltransferase, or an extract of the culture. The present sequence is a protein sequence, for alpha-2,3/alpa-2,8-stalyltransferase from Pasteurella multocida

Sequence 303 AA;

Query Match            44.0%; Score 688.5; DB 6; Length 303;  
Best Local Similarity   45.8%; Pred: No. 6.7e-60;  
Matches 135; Conservative 58; Mismatches 95; Indels 7; Gaps 5

Qy     2   KKVIIAGNGPSLKEIDYSRLENDFDVFRCNQFYFEDKYLGKKCKTVFTPNFFFEQQYT  
Db       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
        11 KAVIVAGNGESLSQIDRYELLPKNYDVFRCNQFYFEERYFLGNKIRKAVFTPGVLEQQYT 70

Qy     62 LKHILIQEYTELIIMCSNYYQAHLNENFKVTFDYDPDAHLGVD-PFKQLKFKNAYPK 120  
Db       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
        71 LYHLKRNNIEYVDNVIISSFNHPTVDLEKSQK-IQALEFIDVINGYEKVLSKLTAFDYLR 129

Qy     121 PHEIYNQRITSGVMCAVALGYKETLYLSGIDEYO-NGSSYAFTDKOENLLKLAPDFK 179  
Db       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
        130 YKEYENQRITSGVMCAVALMGYTDIYLTGIDFYQASEENYAFDNKKPNIIRLLPDFR 189

Qy     180 NDRSHYIGHSKNTDIKALFPLEKTYSIKLYCLCPNSLLANFTEL--APNLNSNFIOEK 236  
Db       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
        190 KEKTLFSVHSKDIDLEALSFLQOHYHNIFYSIPMSPLSKHPPIPTVEDDCETTFVAPLK 249

Qy     237 NNVTYKDLIPSEAYGPSKNINFKIKIKENVYVKLIKDLLRLSPDIKHYPKGX 291  
Db       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
        250 ENYINDILLPHFYVEYKLTIVS-KKSRFHSNLIVRLIRDLLKLPSALKHYLKEK 303

RESULT 15

ID ADNI19559 standard; protein; 958 AA.  
XX AC ADNI19559;  
XX DT 02-DEC-2004 (first entry)  
XX DE Bacterial polypeptide #2212.  
KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactonannan;  
KW bacterial polypeptide.  
OS Bacteria.  
XX PA (CAOY/) CAO Y.  
XX PA (HINK/) HINKLE G J.  
XX PA (SLAT/) SLATER S C.  
XX PA (CHEN/) CHEN X.  
XX PA (GOLD/) GOLDMAN B S.

US2003233675-A1.  
18-DEC-2003.  
20-FEB-2003; 2003US-00369493.  
21-FEB-2002; 2002US-0360039P.





GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 14:55:46 ; Search time 5050.55 Seconds  
(without alignments)  
11053.473 Million cell updates/sec

Title: US-10-734-719-11  
Perfect score: 873  
Sequence: 1 atgaaaaagtattattgc.....agcattatttcaagagaaaa 873

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues  
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_env.\*  
2: gb\_pat.\*  
3: gb\_ph.\*  
4: gb\_pl.\*  
5: gb\_pr.\*  
6: gb\_ro.\*  
7: gb\_sts.\*  
8: gb\_sy.\*  
9: gb\_un.\*  
10: gb\_vi.\*  
11: gb\_ov.\*  
12: gb\_htg.\*  
13: gb\_in.\*  
14: gb\_om.\*  
15: gb\_ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	873	100.0	AX934434 Sequence
2	873	100.0	AR481786 Sequence
3	873	100.0	AR527385 Sequence
4	873	100.0	AR609664 Sequence
5	873	100.0	AR689942 Sequence
6	873	100.0	AR691838 Sequence
7	873	100.0	AF215659 Campyloba
8	868.2	99.5	AF305571 Campyloba
9	866.6	99.3	AX934436 Sequence
10	866.6	99.3	AR481787 Sequence
11	866.6	99.3	AR527386 Sequence
12	866.6	99.3	AR609665 Sequence
13	866.6	99.3	AR689943 Sequence
14	866.6	99.3	AR691839 Sequence
15	866.6	99.3	AX934431 Sequence
16	866.6	99.3	BD249794 Campyloba
17	866.6	99.3	AR271703 Sequence
18	866.6	99.3	AR481785 Sequence

19	866.6	99.3	876	2	AR527384	AR527384 Sequence
20	866.6	99.3	876	2	AR609663	AR609663 Sequence
21	866.6	99.3	876	2	AR689941	AR689941 Sequence
22	866.6	99.3	876	2	AR691837	AR691837 Sequence
23	866.6	99.3	6047	15	AY661458	AY661458 Campyloba
24	866.6	99.3	12390	15	AF401528	AF401528 Campyloba
25	866.6	99.3	24425	15	AY422197	AY422197 Campyloba
26	866.6	99.3	24437	15	AF167344	AF167344 Campyloba
27	849.2	97.3	1303	15	AY681242	AY681242 Campyloba
28	847.4	97.1	876	2	AX934427	AX934427 Sequence
29	847.4	97.1	876	2	BD249792	BD249792 Campyloba
30	847.4	97.1	876	2	AR271701	AR271701 Sequence
31	847.4	97.1	876	2	AR481783	AR481783 Sequence
32	847.4	97.1	876	2	AR527382	AR527382 Sequence
33	847.4	97.1	876	2	AR609661	AR609661 Sequence
34	847.4	97.1	876	2	AR689939	AR689939 Sequence
35	847.4	97.1	876	2	AR691835	AR691835 Sequence
36	845.8	96.9	11442	15	AF400048	AF400048 Campyloba
37	836.2	95.8	876	15	AF216647	AF216647 Campyloba
38	836.2	95.8	6047	15	AY644679	AY644679 Campyloba
39	836.2	95.8	11474	2	AX934424	AX934424 Sequence
40	836.2	95.8	11474	2	BD249790	BD249790 Campyloba
41	836.2	95.8	11474	2	AR271699	AR271699 Sequence
42	836.2	95.8	11474	2	AR481781	AR481781 Sequence
43	836.2	95.8	11474	2	AR527380	AR527380 Sequence
44	836.2	95.8	11474	2	AR609659	AR609659 Sequence
45	836.2	95.8	11474	2	AR689937	AR689937 Sequence

ALIGNMENTS

RESULT 1  
AX934434  
LOCUS AX934434 873 bp DNA linear PAT 05-JAN-2004  
DEFINITION Sequence 11 from Patent WO02074942.  
ACCESSION AX934434  
VERSION AX934434.1 GI:40641680  
KEYWORDS Campylobacter jejuni  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE 1  
AUTHORS Gilbert M. and Wakarchuk W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics  
JOURNAL Patent: WO 02074942-A 11 26-SEP-2002;  
National Research Council of Canada (CA)  
FEATURES  
Location/Qualifiers  
1..873  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1..873  
/note="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
jejuni O:4"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04170.1"  
/db\_xref="GI:40641681"  
translation="MKKVIAGNSLSKIDISRLNDFVFRNQYFEDKYKYLGGK  
CAVAYTGFQFFQYTLKHLQNEIYELIMCSNNOHLENENFVKFYDFPDA  
HLGYDFKQLKEFNAYFKFHEIYFNQRTSGVMCAVAIALGYKIDFYDQNGS  
SYAFDTQENLKLPADPKNDKSHVIGHSKNTDIKALEFKYKIKLYCLCPNSLLA  
NFIELAPLNSNFIQEKNNYTKDILIPSEAVGKFSKNINFKKIKIKENVYKLIKD  
LLRLPSDIKHYFKGK"

ORIGIN

Query Match 100.0%; Score 873; DB 2; Length 873;  
Best Local Similarity 100.0%; Pred. No. 1.3e-245;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGAAAAAGTTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
|
|
|
Db 1 ATGAAAAAGTTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
|
|
|
Qy 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
|
|
|
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
|
|
|
Qy 121 CTGCTGTAATGGAATGATTTTACACCCCTGTTCTTTGAGCAATACTAC 180
|
|
|
Db 121 CTGCTGTAATGGAATGATTTTACACCCCTGTTCTTTGAGCAATACTAC 180
|
|
|
Qy 181 ACTTTAAACATTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240
|
|
|
Db 181 ACTTTAAACATTTAAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240
|
|
|
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTATTCCT 300
|
|
|
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTATTCCT 300
|
|
|
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360
|
|
|
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360
|
|
|
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
|
|
|
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
|
|
|
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
|
|
|
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
|
|
|
Qy 481 TCTTATGCTTTTGATACCAACAGAAATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 540
|
|
|
Db 481 TCTTATGCTTTTGATACCAACAGAAATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 540
|
|
|
Qy 541 GATGCTCATCTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
|
|
|
Db 541 GATGCTCATCTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
|
|
|
Qy 601 GAAAAAATTTCAAAAATAAAATCTATATGCTTATGCTCTAACAGTCTTTTACGAAATTTT 660
|
|
|
Db 601 GAAAAAATTTCAAAAATAAAATCTATATGCTTATGCTCTAACAGTCTTTTACGAAATTTT 660
|
|
|
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATATACAGAAAAATAAATACTACACT 720
|
|
|
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATATACAGAAAAATAAATACTACACT 720
|
|
|
Qy 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAATTTAATTTT 780
|
|
|
Db 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAATTTAATTTT 780
|
|
|
Qy 781 AAAAAAATAAAAAATTAAGAAATTTTATTTACAGTTGATTAAGATTTAATTAAGATTA 840
|
|
|
Db 781 AAAAAAATAAAAAATTAAGAAATTTTATTTACAGTTGATTAAGATTTAATTAAGATTA 840
|
|
|
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
|
|
|
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
|
|
|
```

```
RESULT 2
AR481786 LOCUS AR481786 873 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 11 from patent US 6699705.
ACCESSION AR481786
VERSION AR481786.1 GI:47243421
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 873)
```

AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics  
JOURNAL Patent: US 6699705-A 11 02-MAR-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source  
1. .873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 873; DB 2; Length 873;  
Best Local Similarity 100.0%; Pred. No. 1.3e-245;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
|  
|  
|  
Db 1 ATGAAAAAGTTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
|  
|  
|  
Qy 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
|  
|  
|  
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
|  
|  
|  
Qy 121 CTGCTGTAATGGAATGATTTTACACCCCTGTTCTTTGAGCAATACTAC 180  
|  
|  
|  
Db 121 CTGCTGTAATGGAATGATTTTACACCCCTGTTCTTTGAGCAATACTAC 180  
|  
|  
|  
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240  
|  
|  
|  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAAATATGTTCTAAT 240  
|  
|  
|  
Qy 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTATTCCT 300  
|  
|  
|  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTATTCCT 300  
|  
|  
|  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 360  
|  
|  
|  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360  
|  
|  
|  
Qy 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420  
|  
|  
|  
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420  
|  
|  
|  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
|  
|  
|  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
|  
|  
|  
Qy 481 TCTTATGCTTTTGATACCAACAGAAATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 540  
|  
|  
|  
Db 481 TCTTATGCTTTTGATACCAACAGAAATTTTAAACAACTTAAAGAAATTTAATGCTTTTAA 540  
|  
|  
|  
Qy 541 GATGCTCATCTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600  
|  
|  
|  
Db 541 GATGCTCATCTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600  
|  
|  
|  
Qy 601 GAAAAAATTTCAAAAATAAAATCTATATGCTTATGCTCTAACAGTCTTTTACGAAATTTT 660  
|  
|  
|  
Db 601 GAAAAAATTTCAAAAATAAAATCTATATGCTTATGCTCTAACAGTCTTTTACGAAATTTT 660  
|  
|  
|  
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATATACAGAAAAATAAATACTACACT 720  
|  
|  
|  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATATACAGAAAAATAAATACTACACT 720  
|  
|  
|  
Qy 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAATTTAATTTT 780  
|  
|  
|  
Db 721 AAAGATATATCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAATTTAATTTT 780  
|  
|  
|  
Qy 781 AAAAAAATAAAAAATTAAGAAATTTTATTTACAGTTGATTAAGATTTAATTAAGATTA 840  
|  
|  
|  
Db 781 AAAAAAATAAAAAATTAAGAAATTTTATTTACAGTTGATTAAGATTTAATTAAGATTA 840  
|  
|  
|  
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
|  
|  
|  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
|  
|  
|

Db	841	CCTAGTGATATAAAGCATTATTTTCAAAGGAAAA	873	DNA	linear	PAT 08-OCT-2004
RESULT 3						
LOCUS	AR527385	Sequence 11 from patent US 6723545.	873 bp	DNA	linear	PAT 08-OCT-2004
DEFINITION	AR527385	Sequence 11 from patent US 6723545.	873 bp	DNA	linear	PAT 08-OCT-2004
ACCESSION	AR527385	Sequence 11 from patent US 6723545.	873 bp	DNA	linear	PAT 08-OCT-2004
VERSION	AR527385.1	GI:53914406	873 bp	DNA	linear	PAT 08-OCT-2004
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 873)					
AUTHORS	Gilbert, M. and Wakarchuk, W. W.					
TITLE	Polypeptides having .beta.-1,4-GalNAc transferase activity					
JOURNAL	Patent: US 6723545-A 11 20-APR-2004;					
	National Research Council of Canada; Ottawa;					
	CAX;					
FEATURES	Location/Qualifiers					
source	1..873					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						
	Query Match	100.0%;	Score 873;	DB 2;	Length 873;	
	Best Local Similarity	100.0%;	Pred. No. 1.3e-245;			
	Matches 873;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAGAAATGATTATTC	60			
Db	1	ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAGAAATGATTATTC	60			
Qy	61	CTACCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT	120			
Db	61	CTACCAAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT	120			
Qy	121	CTTGTTAAAAATGCAAAAGCAGTGTGTTTACACCCCTGGTTTCTTTGAGCAATCTAC	180			
Db	121	CTTGTTAAAAATGCAAAAGCAGTGTGTTTACACCCCTGGTTTCTTTGAGCAATCTAC	180			
Qy	181	ACTTTAAAAATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	240			
Db	181	ACTTTAAAAATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	240			
Qy	241	TACAAACAACTTATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	300			
Db	241	TACAAACAACTTATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	300			
Qy	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTT	360			
Db	301	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTT	360			
Qy	361	TTTACGAAATTTTATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	420			
Db	361	TTTACGAAATTTTATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA	420			
Qy	421	ATAGCCCTAGGATACAAAGAAATTTTCTTTTCGGGAATTTGATTTTAA	480			
Db	421	ATAGCCCTAGGATACAAAGAAATTTTCTTTTCGGGAATTTGATTTTAA	480			
Qy	481	TCATTGCTTTTGATACCAACAAAGAAATTTCTTTTAAACCTAGCCCTGATTTT	540			
Db	481	TCATTGCTTTTGATACCAACAAAGAAATTTCTTTTAAACCTAGCCCTGATTTT	540			
Qy	541	GATCGCTCACATATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTT	600			
Db	541	GATCGCTCACATATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTT	600			
Qy	601	GAAGAACTTACAAATATACTATATGCTTATGCTTAAAGCTTTTAAAGCTTT	660			
Db	601	GAAGAACTTACAAATATACTATATGCTTATGCTTAAAGCTTTTAAAGCTTT	660			

```
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY      481  TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540
Db      481  TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540
QY      541  GATCGCTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db      541  GATCGCTACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY      601  GAAAGAACTTACAAATAAACTATATGCTTTATGCTTATGCTTAAACAGTCTTTAGCAATTTT 660
Db      601  GAAAGAACTTACAAATAAACTATATGCTTTATGCTTATGCTTAAACAGTCTTTAGCAATTTT 660
QY      661  ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAAGAAAAATAACTACACT 720
Db      661  ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAAGAAAAATAACTACACT 720
QY      721  AAAGATATACCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Db      721  AAAGATATACCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
QY      781  AAAAAATAAATTAAGAAATCTTTTATTAACAGTTTATTAACAGTCTTATTAAGATTA 840
Db      781  AAAAAATAAATTAAGAAATCTTTTATTAACAGTTTATTAACAGTCTTATTAAGATTA 840
QY      841  CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873
Db      841  CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873

RESULT 5
LOCUS   AR689942                      873 bp      DNA      linear      PAT 12-SEP-2005
DEFINITION   Sequence 11 from patent US 6905867.
ACCESSION   AR689942
VERSION     AR689942.1  GI:74471950
KEYWORDS    'Unknown.'
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 873)
AUTHORS     Gilbert, M. and Wakarchuk, W.W.
TITLE       Nucleic acids encoding polypeptides with .beta.1,3-galactosyl
            transferase activity
JOURNAL     Patent: US 6905867-A 11 14-JUN-2005;
            National Research Council of Canada; Ottawa;
            CAX;
FEATURES             Location/Qualifiers
     source           1..873
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 873; DB 2; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.3e-245;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGAAAAAGTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60
Db      1  ATGAAAAAGTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTCAGG 60
QY      61  CTACCAATGATTTTATGATGTTATAGATCTATCAATTTTATTTGAGATAAATACTAT 120
Db      61  CTACCAATGATTTTATGATGTTATAGATGTTATCAATTTTATTTGAGATAAATACTAT 120
QY      121  CTTGTAAAAAATGCAAGACAGTGTTTTACACCCCTGGTTTCTTCTTTGAGCAATACTAC 180
Db      121  CTTGTAAAAAATGCAAGACAGTGTTTTACACCCCTGGTTTCTTCTTTGAGCAATACTAC 180
QY      181  ACTTTAAACATTTAATCGAAATCAAGAAATATGACCGAACTAATTAATGTGTTCTAAT 240
Db      181  ACTTTAAACATTTAATCGAAATCAAGAAATATGACCGAACTAATTAATGTGTTCTAAT 240
```

```
Db      181  ACTTTAAACATTTAATCGAAATCAAGAAATATGACCGAACTAATTAATGTGTTCTAAT 240
QY      241  TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTCCCT 300
Db      241  TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTCCCT 300
QY      301  GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAA 360
Db      301  GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTAAA 360
QY      361  TTTTCAGGAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
Db      361  TTTTCAGGAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420
QY      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCA 480
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCA 480
QY      481  TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db      481  TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY      541  GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Db      541  GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY      601  GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660
Db      601  GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660
QY      661  ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCAATTTTATCATACAAAGAAAAATACTACACT 720
Db      661  ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCAATTTTATCATACAAAGAAAAATACTACACT 720
QY      721  AAAGATATACCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
Db      721  AAAGATATACCTATACCTCTTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
QY      781  AAAAAATAAATTAAGAAATCTTTTATTAACAGTTTATTAACAGTCTTATTAAGATTA 840
Db      781  AAAAAATAAATTAAGAAATCTTTTATTAACAGTTTATTAACAGTCTTATTAAGATTA 840
QY      841  CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873
Db      841  CCTAGTGATATAAAGCAATTTTCAAGGAAAA 873

RESULT 6
LOCUS   AR691838                      873 bp      DNA      linear      PAT 13-SEP-2005
DEFINITION   Sequence 11 from patent US 6911337.
ACCESSION   AR691838
VERSION     AR691838.1  GI:74479549
KEYWORDS    'Unknown.'
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 873)
AUTHORS     Gilbert, M. and Wakarchuk, W.W.
TITLE       Nucleic acids encoding .beta.-1,4-GalNAC transferase
JOURNAL     Patent: US 6911337-A 11 28-JUN-2005;
            National Research Council of Canada; Ottawa;
            CAX;
FEATURES             Location/Qualifiers
     source           1..873
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 873; DB 2; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.3e-245;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG 60
Db 1 ATGAAAAAGTTATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTATCAAGG 60
QY 61 CTACCAAAATGATTTGATGATTAATAGATGAATCAATTTTATTTGAAGATAAATACAT 120
Db 61 CTACCAAAATGATTTGATGATTAATAGATGAATCAATTTTATTTGAAGATAAATACAT 120
QY 121 CTTGGTAAAAATGCAAGAGCTGTTTACACCCCTGGTTTCTCTTTGAGCAATATAC 180
Db 121 CTTGGTAAAAATGCAAGAGCTGTTTACACCCCTGGTTTCTCTTTGAGCAATATAC 180
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGCACTAATTAATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGCACTAATTAATGTTCTTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCC 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTAATTTTCC 300
QY 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 361 TTTACGAAATTTAATCAATCAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420
Db 361 TTTACGAAATTTAATCAATCAAGAAATACCTCAGGGGTCTATATGTGCGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGGAATTTGATTTTATCAAAATGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGGAATTTGATTTTATCAAAATGGTCA 480
QY 481 TCTTATGCTTTTATACCAAAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTATACCAAAACAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
QY 541 GATCCTCACAATATATCGGATATGATTAATAATACAGATATATAAGCTTTAGAAATTTCA 600
Db 541 GATCCTCACAATATATCGGATATGATTAATAATACAGATATATAAGCTTTAGAAATTTCA 600
QY 601 GAAAAAATTTACAAATAAATAATTTATGCTTATGCTTAAACAGTCTTTTACAAATTTT 660
Db 601 GAAAAAATTTACAAATAAATAATTTATGCTTATGCTTAAACAGTCTTTTACAAATTTT 660
QY 661 ATAGAACTAGCCCAATTTAAATTTAAATTTTATCATAACAGAAAAAATAACTACAT 720
Db 661 ATAGAACTAGCCCAATTTAAATTTAAATTTTATCATAACAGAAAAAATAACTACAT 720
QY 721 AAAGATATATCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTT 780
Db 721 AAAGATATATCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATTTAATTT 780
QY 781 AAAAAATTAATAATTAAGAAATGTTTATTAAGAGTTGATTAAGATCTATAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATGTTTATTAAGAGTTGATTAAGATCTATAGATTA 840
QY 841 CCTAGTGATATAAGCATTTTTCACAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTTTCACAGGAAAA 873
```

```
RESULT 7
LOCUS AF215659 11474 bp DNA linear BCT 02-JAN-2001
DEFINITION Campylobacter jejuni lipooligosaccharide biosynthesis locus,
partial sequence.
ACCESSION AF215659
VERSION AF215659.1 GI:12004275
KEYWORDS
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

FEATURES  
source

source

source

CDS

CDS

CDS

CDS

gene

1 (bases 1 to 11474)  
Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.  
The LOS biosynthesis locus of Campylobacter jejuni O:4  
Unpublished

2 (bases 1 to 11474)  
Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.  
Direct Submission  
Submitted (12-DEC-1999) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A  
0R6, Canada

Location/Qualifiers

1. 11474

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strain="ATCC 43432"

/serotype="O:4"

/db\_xref="ATCC:43432"

/db\_xref="taxon:197"

<1. .360

/note="ORF-01; similar to RfaC from Helicobacter pylori"

/codon start=1

/transl table=11

/product="heptosyltransferase I"

/protein\_id="AAG43973.1"

/db\_xref="GI:12004276"

/translation="KEYFAKEVNLGSGIDETHIELAPKFSLEELMAFTQWDLIIGN  
DSGPTHLAFLAKASITIFGATPSYRNAFQTHINKIIDAGKIQNAKHIDKSDFCITR  
IREEDIFKLAKGLLNEK"

350. .1237

/note="ORF-02; similar to WaaM from Helicobacter pylori"

/codon start=1

/transl table=11

/product="acylttransferase"

/protein\_id="AAG43974.1"

/db\_xref="GI:12004277"

/translation="MKNSDRIYLSYILKFFVTMPECILHFLALIVARITFYLNKK  
HRKIIDINQICFPQYTKERDKSLKIYENFAQGDICLQNTTKEIKLVNFIN  
ENLLIDALKRPIITTAHYGNWEILSLAAYKAGISIVGKLKSEIYVILSONR  
TOPDIELDKGGIROMLSALKGRALGILTDODCVKESVRLKFNKEVYOMGASL  
IAORSNALLIPVYAYKEGKFCIEFFAKDSQNASLEELTYLQAQSCSEMIKKRPWEY  
FFHRRPASNYEEIYKAK"

1234. .2490

/note="ORF-03; similar to LgtF from Neisseria"

/meningitidis"

/codon start=1

/transl table=11

/product="glycosyltransferase"

/protein\_id="AAG43975.1"

/db\_xref="GI:12004278"

/translation="MNLKQISVIIIVKNAQOTLLECLNSLKDPEIILNNESIDNTL  
NIANKPKDPANLYIYHNAFIGALKNLALSAYAKNDWILSIDADEVLENCIKELKN  
LKQENIIALSRKNLYKGEWIKACGWPDYVIRIFNKNTRFRNDNLVHESLIVPSNA  
KKIKNGKLKHSYKIDSHLIDKQYSSSLWAKONHKKGVLANRAFAWTFPRNYF  
LKNGLYGYKGFIIISVCSALGTFKMYLQKPEKQKPCALIIITYNKRLKLVLD  
SVKNAFLPNEVLIIADDSKEDTARLIEYKQDPCPLKHIWQDEGFKLSKNKTI  
KNADSVIIVIDGMILEKDFIKHELEFAOKFLQGSRVILNKSEELINKDDYRI  
IFNKDPKSKNSFLAKIFYLSKKR"

2786. .3955

/note="ORF-04; local similarity (N-terminus) to Cps14J  
from Streptococcus pneumoniae"

/codon start=1

/transl table=11

/product="glycosyltransferase"

/protein\_id="AAG43976.1"

/db\_xref="GI:12004279"

/translation="MKKIGWIPYINVEKYRECLDSVINQVNTYNLEIILVNDGSTDE  
HSLNTAKVTLKDKRTTFDKNGGLSSARNIGIEYPSGKVLKNKTOHKIENSLIEF  
OLDGNPNYIYKAKSSQAENKDLTNTYPSIDYITFLDSNWKLVNCEVIRIM  
XNDVLFPHDCTIEDYINKNKHKTRMEIFDFKKECLITPEYANRALSVGSRDISFG  
WNGMIDTFNLKQIKLFIPIINEDIHGILIFASANKIYVLSQVLCRLRANSIN  
HDKKITKANVSEYFKDIYETFGENAKENYKAAARVITALKLIEFFKQOKNENALA  
IKETFLPCYAKKALMIKKFKKQPLNLKEQLVLIKPFITQTKLPYDINKFWQIKNI"  
4025. .5068

CDS	<pre>/gene="cgtA" 4025..5068 /gene="cgtA" /note="glycosyltransferase; similar to CgtA from Campylobacter jejuni OH4384 encoded by GenBank Accession Number AF130984" /codon_start=1 /transl_table=11 /product="beta-1,4-N-acetylglactosaminyltransferase" /protein_id="AAG43977.1" /db_xref="GI:12004280"</pre>	<pre>/product="sialic acid biosynthesis protein" /protein_id="AAG43981.1" /db_xref="GI:12004284" /translation="MYRQNSSEFELYIFATGMHLSKNFGYTVKELYKMGFKNIYERFI NYDKYQTDKALATIDGFSRYVNEKLKEDLIVVHGDRIEPLAAAIWALNNILVAHIE GGEISGTIDDSLRHAIKSLAHILVNDFAKRLMLQGEDEKSIIFIGSPDLDELNDN KLSLSEAKYQIDINYENYALLMFHVTEITSIKQADNLKALIGOSNNKNIYIYPPNN DLGFELILQSYBELKNRPFKLFPSPRFEYFITLLKNADFIIGNSSCIKEALYKTA LIGVSRQNGRLNGENTLKVANSDEILKAIINTIHKQDLFSAKLEILDLSKLFPEYL LSGFEFKLNTQKVKDIK" 9076..9741 /note="ORF-10; similar to NeuA from Haemophilus ducreyi" /codon_start=1 /transl_table=11 /product="CMP-sialic acid synthetase" /protein_id="AAG43982.1" /db_xref="GI:12004285" /translation="MSLAIIPARGSGKIGKNKLVLLNNKPLIYTIKALNTKSIK VWSSDSDEILNYSKSNVDILKRPIISAQDDTSDKVLHLKFKYDYEDVVFLQPT SPLRTHIHIDEAFNLKSNANALISVSECNDKILKAFVNDYGDLAGLNDYEPFIP ROKLPKTYMSNGAIYILKIEFLNPNFLQNKTKHFLMDESSLDIDCLEDLKKVEGI WKK" 9729..10562 /note="ORF-11; similar to putative ORF from Rhodobacter capsulatus encoded by GenBank Accession Number AF010496" /codon_start=1 /transl_table=11 /product="acetyltransferase" /protein_id="AAG43983.1" /db_xref="GI:12004286" /translation="MEKITLCKNKNILLLKOYNIYTKYIENRRFSRLKTKDFTTI PLENNQLSVAGLIEEYAFKSNILHEMGFSFSGSLPHYAKVGRVCSIAAGVSM FNQHPMRISTASTYETNSFINNACONHINKTFFPIVHNPPSSITHLIIODDVMII CKVLKQIGITLGTCTGIVGQAVTKDPPYVAIVAGIPAKIILKYRKNKTIERLLKIQ WKKHFPAPFYDIDLNLKINQYLDLLEEKIKKSIYYNPNKLYFRDILELKSXKIFNL F"</pre>
CDS	<pre>/complement(5054..5959) /gene="cgtB" /note="cgtB" complement(5054..5959) /gene="cgtB" /note="glycosyltransferase; similar to CgtB from Campylobacter jejuni OH4384 encoded by GenBank Accession Number AF130984" /codon_start=1 /transl_table=11 /product="beta-1,3-galactosyltransferase" /protein_id="AAG43978.1" /db_xref="GI:12004281" /translation="MFKISILPTYNVEQYIARIEASCINQTFKDIIEIIVVDDCGNDN SINIAYESKDKRIKIHNEKNLGLRAREYGVKANSPIIMFLDDPDYLELNACE CIKLDQDEVDLVFNAIVESNIVSYKGFDFNSGFYSKFEVKKI IAKNKLWTMNG KILRKLYLBAFASLEKQDVKNNAEDVLLYFPMLSQAQKIAVMNCLXHVPPNNS ICNTKRVILQNNIQELQVLNLYRQNYLNNKYCSVLVLYKILLYIQIYKIKTKLM VTLAKNIITLILKILFYKKFLKQC" 6048..6923 /gene="cst-II" 6048..6923 /gene="cst-II" /note="sialyltransferase; activity determined experimentally" /codon_start=1 /transl_table=11 /product="alpha-2,3-sialyltransferase" /protein_id="AAG43979.1" /db_xref="GI:12004282" /translation="MKKVI IAGNPSLKEIDYSLPNDFVRCNQPFYFEDKYVLGKK CKAVYTPGFFFEQYTLKHLIQOEYETELIMCSNYNOAHLENENFVKTFDYFPDA HLGYDFPQLKEFNAYFKFHEIYFNQRTITSGVMCAVAIALGYKEIYLSGIDFYONGS SYAFDTKOENLLKLAPDKNDRSHVIGHSKNTDIIKALEFLEKTKIKLYCICPNSLLA NFLELAPLNSNFIIOENKNTKDILIPSEAYGKFSKNINFKIKIKIKENVYIKLIKID LLRLPDIKHYFKGK" 6924..7964 /note="ORF-08; similar to SiaC from Neisseria meningitidis" /codon_start=1 /transl_table=11 /product="sialic acid synthase" /protein_id="AAG43980.1" /db_xref="GI:12004283" /translation="MKKIKIQNIISSEKAPLVPEIIGINHNGSLELAKIMVDAFSA GAKIITHQTHIVDEMSEKAAKVIIPGNAKISIVEIMOKCALNYKDELAKLEYTEKLG VYLSTPSRAGVNRLEMGVSFAKIGSGECNNYPLIKHIAAFKPKMIVSTGMNIESI LKVILLNDNEIPVLMNLNLYPTPHNLVRLNMLKEKFSKCMVGLSDHTHSLG LGVALGACVLERHTDSMRSGPDIVCSMDTQALKELIIOSEQWALMRGNESKAA KQEQVTDIFAFASVVSIKDIKKGEVLSDMNIWVRPGLGGISASEFENILGKALRDI ENDTOLSYEDFT" 8021..9079 /note="ORF-09; similar to SiaA from Neisseria meningitidis" /codon_start=1 /transl_table=11</pre>	<pre>complement(10554..11366) /note="ORF-12; similar to putative ORF HI0868 from Haemophilus influenzae encoded by GenBank Accession Number U32768" /codon_start=1 /transl_table=11 /product="glycosyltransferase" /protein_id="AAG43984.1" Query Match 100.0%; Score 873; DB 15; Length 11474; Best Local Similarity 100.0%; Pred. No. 7.4e-246; Mismatches 0; Indels 0; Gaps 0; Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60 Db 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 6107 QY 61 CTACCAAAATGATTTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120 Db 6108 CTACCAAAATGATTTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 6167 QY 121 CTGTGTAATAAATGCAAGACAGTGTTTACACCCCTGGTTCTTCTTTGAGCAATACTAC 180 Db 6168 CTGTGTAATAAATGCAAGACAGTGTTTACACCCCTGGTTCTTCTTTGAGCAATACTAC 6227 QY 181 ACTTTAAACATTTTAAATCAAAATCAAGAAATATGAGCCGAACCTAATTTATGTTCTTAAT 240 Db 6228 ACTTTAAACATTTTAAATCAAAATCAAGAAATATGAGCCGAACCTAATTTATGTTCTTAAT 6287 QY 241 TACACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACTTTTTCAGGATTATTTTCCT 300 Db 6288 TACACCAAGCTCATCTAGAAATGAAAATTTTGTAAAAACTTTTTCAGGATTATTTTCCT 6347 QY 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360 Db 6348 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 6407</pre>

```
QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTADCC 420
Db 6408 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTADCC 6467
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAATTTGATTTTTATCAAAATGGGTCA 480
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGCGAATTTGATTTTTATCAAAATGGGTCA 6527
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Db 6528 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6587
QY 541 GATCGCTCACACTATATCGGACATAGTAAATACAGATATATAAAGCTTTAGAAATTTCTA 600
Db 6588 GATCGCTCACACTATATCGGACATAGTAAATACAGATATATAAAGCTTTAGAAATTTCTA 6647
QY 601 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTAAACAGTCTTTTAGCAATTTT 660
Db 6648 GAAAAAATTTACAAATAAAATCTATATGCTTATGCTTAAACAGTCTTTTAGCAATTTT 6707
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 720
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 6767
QY 721 AAAGATATACATACACTTCTAGTGAGGCTTATGAAATTTTCAAAAAATTAATTTT 780
Db 6768 AAAGATATACATACACTTCTAGTGAGGCTTATGAAATTTTCAAAAAATTAATTTT 6827
QY 781 AAAAAAATAAAATTTAAAGAAATGTTTATTAACAAGTTCATAAAAGATCTATTAAAGATTA 840
Db 6828 AAAAAAATAAAATTTAAAGAAATGTTTATTAACAAGTTCATAAAAGATCTATTAAAGATTA 6887
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
Db 6888 CCTAGTGATATAAGCAATTTTCAAGGAAAA 6920

RESULT 8
AF305571
LOCUS
DEFINITION
Campylobacter jejuni strain 81-176 sialyl transferase (cst), sialic
acid synthase (neuB), NeuC, N-acetylglucosyltransferase, and
CMP-NeuNac synthetase (neuA) genes, complete cds.
ACCESSION
AF305571
VERSION
AF305571.1 GI:15824490
KEYWORDS
Campylobacter jejuni subsp. jejuni 81-176
Campylobacter jejuni subsp. jejuni 81-176
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
SOURCE
1 (bases 1 to 4749)
Guerry, P., Symanski, C.M., Prendergast, M.M., Hickey, T.E.,
Ewing, C.P., Pattarini, D.L. and Moran, A.P.
Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide
affects ganglioside mimicry and invasiveness in vitro
Infect. Immun. 70 (2), 787-793 (2002)
JOURNAL
PUBMED
11796612
REFERENCE
2 (bases 1 to 4749)
Guerry, P., Symanski, C.M., Prendergast, M.M., Hickey, T.E.,
Ewing, C.P., Pattarini, D.L. and Moran, A.P.
Direct Submission
Submitted (14-SEP-2000) Enteric Diseases, Naval Medical Research
Center, 503 Robert Grant Ave, Silver Spring, MD 20910, USA
FEATURES
Location/Qualifiers
1..4749
/organism="Campylobacter jejuni subsp. jejuni 81-176"
/mol_type="genomic DNA"
/strain="81-176"
/db_xref="taxon:354242"
89..964
/gene="cst"
89..964
/gene="cst"

gene
CDS
ORIGIN
```

```
/note="Cst"
/codon_start=1
/transl_table=11
/product="sialyl transferase"
/protein_id="AAL09368.1"
/db_xref="GI:15824491"
/translation="MKKVIITAGNGPSLKIEDYSLRPNDFVRCNQFVPEFKYILGKK
CKAFVYTFNFFQYITLKHILQOEYETELIMCSNYNQAHLENFVKTFDYDPDA
HLGYDFKQLKFNAYFKPHEIYFNQITSGVMCAVAIALGYKEIYLSGLDFYQNGS
SYAFDTQENLLKLAPDKNDRSHYIGHSKNTDIKALEFLEKTKYIKLYCLCPNLLA
NFTELAPNLNSNFIIOEKNNYTKDILIPSSSEAYGKFSKNINFKKIKIKENVYKLIKD
LLRLPSDILKHYFKGK"
965..2005
/gene="neuB1"
965..2005
/gene="neuB1"
/notes="NeuB1"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAL09369.1"
/db_xref="GI:15824492"
/translation="MKEIKIQNIILISEKAPLVVPEIGINHGSLSLAKIMVDAFSA
GAKILKHQTHIVEDSKAAKVIIPGNAKIYIEMOKCALDYDKDELALKEYTEKGL
VYLSTPFRAGANRLEDMGSAFKIGSGECNNYELIHHAAFKPMIVSTGWNSTESI
KPTVKILLDNEIIPVLMHSTNLPTPHNLVRLNMLEKKEFSQVMGLSDHTNLLAC
LGAVALGACVLERHFTDSMRHSGPDI VCSMDTKALKELIIQSEOMAMKGNESKAA
KOBQVTTIDFAFASVVSVDIKDKGEVLSMDNIWVKRPGLGISAAEFENILGKKALRDI
ENDTQLSYDEPA"
2062..3126
/notes="N-acetylmannosamine synthesis"
/codon_start=1
/transl_table=11
/product="NeuC1"
/protein_id="AAL09370.1"
/db_xref="GI:15824493"
/translation="MYRVQNSSEFELYIFATGMHLSKNFGYTVVKELKNGKFIYEFI
NYDKVYQTKALATTIDGFSRYVYNELKPDILIVVHGDRIEPLAAAVIGALNNILVAHIE
GEISGTIDSLRHAISKLAIHILVNDDEFARRMOLGEDEKSTFIIGSPDLELNDN
KISLSAKKYDINENYALLMEHPVYTTTSIKNOADNLVKALIQSNKNIYIVPNN
DIGFEILQSYEELKNNRPFKPFPSURFEYFILLKNADFIIGNSSCILKEALYKTA
GLVSGRNGRLGNTLKVNSANSDEILKAINTIHKKQDLFSTKLEILDSSSKLPFEYL
LSGEFFKLSTQKVFKDKHA"
3119..4066
/notes="CgtA"
/codon_start=1
/transl_table=11
/product="N-acetylglucosyltransferase"
/protein_id="AAL09371.1"
/db_xref="GI:15824494"
/translation="MLKKIISLYKRYISIKLVLDNEHFIKENKNIYKHKHKGPFDP
EKADVKSPLNPGWFRVKNBAUTLRVLESILPALQRIIAYNDCDDGSELILEFC
KQFNFIAKKYPKVDLENPKNEENKLYSYNWAAGFIPLDWEFKIDVDHYDAKKL
KYSFINDKALCYPRINFILNGNIYVQNSNGFIFGGQDLILKRNSSFIERR
VSKKQWIDPKGLIEELYSEQQVLSQVKILOAPLQWHPALPKYRRNDYQQYLDILS
LEEFQAFHRKSKEAKKIDFAMLKRPVIEQILKKFQGEIK"
4063..4728
/gene="neuA1"
4063..4728
/gene="neuA1"
/notes="NeuA1"
/codon_start=1
/transl_table=11
/product="CMP-NeuNac synthetase"
/protein_id="AAL09372.1"
/db_xref="GI:15824495"
/translation="MSLAIIPARGSGSGIKNNKVLNNKPLIYVITKAALNAKSISK
VVSSDSEILNAYKSONVIDLKRPISLAQDDTTSKVLHLKFKYKDYEDVFLQPT
SPLRTNIHDKAFNLKNSNANALISVSECNDKILKAFVNCNDIGDLAGIENDEYFPNP
ROKLPKTYMSNGAIYILKIKFELNNPSPFLQNKTKHFLMDESSSIDIDCLEDLKKVEQI
WKK"
```



Query Match		99.5%; Score 868.2; DB 15; Length 4749;
Best Local Similarity		99.7%; Pred. No. 2.3e-244;
Matches		870; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy	1	ATGAAAAAGTTATTATTCGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db	89	ATGAAAAAGTTATTATTCGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 148
Qy	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db	149	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 208
Qy	121	CTTGCTTAAATAATGCAAGCAGTGTTTTACACCCCTGGTTCCTCTTTTGACGAATACAC 180
Db	209	CTTGCTTAAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTTGACGAATACAC 268
Qy	181	ACTTTAAACATTTAAATCCAAATCAAGATATGAGCCGGAACCTAAATATGTTCTTAAT 240
Db	269	ACTTTAAACATTTAAATCCAAATCAAGATATGAGCCGGAACCTAAATATGTTCTTAAT 328
Qy	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCTT 300
Db	329	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCTT 388
Qy	301	GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
Db	389	GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 448
Qy	361	TTTACCAAAATTTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGAGTAGCC 420
Db	449	TTTACCAAAATTTTCAATCAAGAAATTAACCTCAGGGTCTATATGTCGAGTAGCC 508
Qy	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db	509	ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 568
Qy	481	TCCTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540
Db	569	TCCTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 628
Qy	541	GATGCTCAGACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
Db	629	GATGCTCAGACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 688
Qy	601	GAATAAACTTACAAAAATAAACTATATTGCTTATGCTTACAGCTCTTTTACGCAAAATTT 660
Db	689	GAATAAACTTACAAAAATAAACTATATTGCTTATGCTTACAGCTCTTTTACGCAAAATTT 748
Qy	661	ATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAGAAAAAAATACTACACT 720
Db	749	ATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAGAAAAAAATACTACACT 808
Qy	721	AAAGATATCTACACTCTAGGAGCTTTATGGAATAATTTTCAAAAAATATTAATTTT 780
Db	809	AAAGATATCTACACTCTAGGAGCTTTATGGAATAATTTTCAAAAAATATTAATTTT 868
Qy	781	AAAAAAATAAAATTAAGAAAAATGTTTATTACAGTTGATAAGATCTATTAGATTA 840
Db	869	AAAAAAATAAAATTAAGAAAAATGTTTATTACAGTTGATAAGATCTATTAGATTA 928
Qy	841	CCTAGTGATATAAGCAATTTTTCAAAAGGAAA 873
Db	929	CCTAGTGATATAAGCAATTTTTCAAAAGGAAA 961
RESULT 9		
AX934436		873 bp DNA linear PAT 05-JAN-2004
LOCUS		
DEFINITION		Sequence 13 from Patent WO02074942.
ACCESSION		AX934436
VERSION		AX934436.1 GI:40641682
KEYWORDS		Campylobacter jejuni
SOURCE		
ORGANISM		Campylobacter jejuni
REFERENCE		Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.
AUTHORS		1 Gilbert, M. and Wakarchuk, W.W.
TITLE		Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics
JOURNAL		Patent: WO 02074942-A 13 26-SEP-2002;
FEATURES		National Research Council of Canada (CA)
source		Location/Qualifiers
1..873		
/organism="Campylobacter jejuni"		
/mol_type="unassigned DNA"		
/db_xref="taxon:197"		
1..->873		
/note="unnamed protein product; Campylobacter		
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.		
jejunii O:36"		
/codon_start=1		
/transl_table=11		
/protein_id="CAP04171.1"		
/db_xref="GI:40641683"		
/translation="MKKVIILAGNPSLKEIDYSLPNDPDPVRCNQFYEDKVLGKK		
CKTVFYTNFFEQIYTLKHLIOEYETELIMCSYNQAHLENENFVKTFDYFDG		
HLGYDFKQLEFNAYFKHEIYFNQITSGVMCAVAIALGYKEIYLSGIDFYQGS		
SYADFTQENLLKLAPDKNDRSHYIGHSKNTDIKALEFLETKYIKLYCLCPNLSLA		
NFIELAPNLNFIIOEKNNVTKDILIPSEAYGFKSNINFKIKIKENVVYKLIKD		
LLRUPSDIKHYFKGK"		
ORIGIN		
Query Match		99.3%; Score 866.6; DB 2; Length 873;
Best Local Similarity		99.5%; Pred. No. 1e-243;
Matches		869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy	1	ATGAAAAAGTTATTATTCGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Db	1	ATGAAAAAGTTATTATTCGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Db	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy	121	CTTGCTTAAATAATGCAAGCAGTGTTTTACACCCCTGGTTCCTCTTTGACGAATATCTAC 180
Db	121	CTTGCTTAAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTCTTTGACGAATATCTAC 180
Qy	181	ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGGAACCTAAATATGTTCTTAAT 240
Db	181	ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGGAACCTAAATATGTTCTTAAT 240
Qy	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCTT 300
Db	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCTT 300
Qy	301	GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Db	301	GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy	361	TTTACGAAATTTTATTCAAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Db	361	TTTACGAAATTTTATTCAAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Qy	421	ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy	481	TCCTATGCTTTTGATACCAAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540
Db	481	TCCTATGCTTTTGATACCAAAATCTTTTAAACCTAGCCCTCGATTTTAAAAAT 540
Qy	541	GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
Db	541	GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600

Qy	361	TTTCAAGAAATTTATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGTGACGTAGCC	420			
Db	361	TTTCAAGAAATTTATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGTGACGTAGCC	420			
Qy	421	ATAGCCCTTAGGATACAAGAAATTTATCTTTTCGGGAATTGATTTTTATCAAAATCGGTCA	480			
Db	421	ATAGCCCTTAGGATACAAGAAATTTATCTTTTCGGGAATTGATTTTTATCAAAATCGGTCA	480			
Qy	481	TCCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540			
Db	481	TCCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540			
Qy	541	GATCGCTCACACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTGAATTTCTTA	600			
Db	541	GATCGCTCACACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTGAATTTCTTA	600			
Qy	601	GAAGAACTTACAAATTAAGAACTATATGCTTATGTCCTAACAGCTTTTACCAATTTT	660			
Db	601	GAAGAACTTACAAATTAAGAACTATATGCTTATGTCCTAACAGCTTTTACCAATTTT	660			
Qy	661	ATAGAACTAGGCCCAATTTTAAATTTCAATTTTATCATACAAGAAAAATAAATACTACT	720			
Db	661	ATAGAACTAGGCCCAATTTTAAATTTCAATTTTATCATACAAGAAAAATAAATACTACT	720			
Qy	721	AAAGATATACTCATACCTCTTAGTGAGCTTATGGAATAATTTTCAAAAAATATTAAATTT	780			
Db	721	AAAGATATACTCATACCTCTTAGTGAGCTTATGGAATAATTTTCAAAAAATATTAAATTT	780			
Qy	781	AAAAAATAAATAATTAAGAAATCTTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA	840			
Db	781	AAAAAATAAATAATTAAGAAATCTTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA	840			
Qy	841	CTTAGTGATATAAGACATTTTTCAAAAGGAAAA	873			
Db	841	CTTAGTGATATAAGACATTTTTCAAAAGGAAAA	873			
RESULT 11						
AR527386						
LOCUS	AR527386	873 bp	DNA linear PAT 08-OCT-2004			
DEFINITION	Sequence 13 from patent US 6723545.					
ACCESSION	AR527386					
VERSION	AR527386.1 GI:53914407					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1. (bases 1 to 873)					
AUTHORS	Gilbert, M. and Wakarchuk, W.W.					
TITLE	Polypeptides having .beta.-1,4-GalNac transferase activity					
JOURNAL	Patent: US 6723545-A 13 20-APR-2004;					
National Research Council of Canada, Ottawa;						
CAX;						
FEATURES						
source	Location/Qualifiers					
	1. .873					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						
Query Match 99.3%; Score 866.6; DB 2; Length 873;						
Best Local Similarity 99.5%; Pred. No. 1e-243;						
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;						
Qy	1	ATCAAAAAAGTTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG	60			
Db	1	ATCAAAAAAGTTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG	60			
Qy	61	CTACCAAAATCATTTTGATGTATTAGATGAATCAATTTTATTTTGAAGATAAAATCTAT	120			
Db	61	CTACCAAAATCATTTTGATGTATTAGATGAATCAATTTTATTTTGAAGATAAAATCTAT	120			
Qy	121	CTTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGTTTCTTTCTTTTGAGCAATCTAC	180			

```
Db 121 CTTGGTAAAAATGCAAAACAGTGTTTTCACACCCCTAAATTTCTTCTTTGGACAATACCTAC 180
Qy 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAAATATGAGACCGGAACCTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAAAACATTTAAATCCAAAATCAAGAAATATGAGACCGGAACCTAAATATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATTAATTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATTAATTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAA 360
Qy 361 TTTTACGAAAATTTTAAATCAAAAGAAATTAATCACTCAGGGGCTATATGTGTGAGTAGGCC 420
Db 361 TTTTACGAAAATTTTAAATCAAAAGAAATTAATCACTCAGGGGCTATATGTGTGAGTAGGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTATAGAAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTATAGAAATTTCTA 600
Qy 601 GAAAAAATCAAAAATTAAGAAATTTTATGCTTATGCTCTAATAGCTTTTATAGCAAAATTTT 660
Db 601 GAAAAAATCAAAAATTAAGAAATTTTATGCTTATGCTCTAATAGCTTTTATAGCAAAATTTT 660
Qy 661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAATTAACACTACACT 720
Db 661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAATTAACACTACACT 720
Qy 721 AAAGATATACCTACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATTAATTAATTTT 780
Db 721 AAAGATATACCTACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATTAATTAATTTT 780
Qy 781 AAAAAAATAAAAATTAAGAAATTTTATTAACAAGTTGATTAAGAAATCTATTAAGATTA 840
Db 781 AAAAAAATAAAAATTAAGAAATTTTATTAACAAGTTGATTAAGAAATCTATTAAGATTA 840
Qy 841 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 873
Db 841 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 873
```

RESULT 12  
AR609665  
LOCUS AR609665 873 bp DNA linear PAT 15-DEC-2004  
DEFINITION Sequence 13 from patent US 6825019.  
ACCESSION AR609665  
VERSION AR609665.1 GI:56664965  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 873)  
AUTHORS Gilbert M. and Wakarchuk, W. W.  
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity  
JOURNAL Patent: US 6825019-A 13 30-NOV-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source Location/Qualifiers  
1..873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

```
Query Match 99.3%; Score 866.6; DB 2; Length 873;  
Best Local Similarity 99.5%; Pred. No. 1e-243;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGCTTATTTATTTGCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Db 1 ATGAAAAAGCTTATTTATTTGCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60  
Qy 61 CTACCAAAATGATTTTGTAGTGTATTAATCAATTTTATTTTCAAGAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGTAGTGTATTAATCAATTTTATTTTCAAGAGATAAATACTAT 120  
Qy 121 CTTGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTGAGCAATACTAC 180  
Db 121 CTTGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180  
Qy 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATATATGTTCTAAT 240  
Db 181 ACTTTAAAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATATATGTTCTAAT 240  
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTAATTTTCC 300  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTAATTTTCC 300  
Qy 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360  
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360  
Qy 361 TTTTACGAAAATTTTATTTCAAAAGAAATTAATCACTCAGGGGCTATATGTGTGAGTAGGCC 420  
Db 361 TTTTACGAAAATTTTATTTCAAAAGAAATTAATCACTCAGGGGCTATATGTGTGAGTAGGCC 420  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTTTATCAAAAATGGGTCA 480  
Qy 481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAAT 540  
Db 481 TCTTATGCTTTTGATACCAACAGAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAAT 540  
Qy 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTATAGAAATTTCTA 600  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTATAGAAATTTCTA 600  
Qy 601 GAAAAAATCAAAAATTAAGAAATTTTATGCTTATGCTCTAATAGCTTTTATAGCAAAATTTT 660  
Db 601 GAAAAAATCAAAAATTAAGAAATTTTATGCTTATGCTCTAATAGCTTTTATAGCAAAATTTT 660  
Qy 661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAATTAACACTACACT 720  
Db 661 ATAGAACTAGCCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAATTAACACTACACT 720  
Qy 721 AAAGATATACCTACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATTAATTAATTTT 780  
Db 721 AAAGATATACCTACACTCTTAGTGAGGCTTATGGAATAATTTTCAAAAATTAATTAATTTT 780  
Qy 781 AAAAAAATAAAAATTAAGAAATTTTATTAACAAGTTGATTAAGAAATCTATTAAGATTA 840  
Db 781 AAAAAAATAAAAATTAAGAAATTTTATTAACAAGTTGATTAAGAAATCTATTAAGATTA 840  
Qy 841 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 873  
Db 841 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 873
```

RESULT 13  
AR689943  
LOCUS AR689943  
DEFINITION Sequence 13 from patent US 6905867.  
ACCESSION AR689943  
VERSION AR689943.1 GI:74471951  
KEYWORDS  
SOURCE Unknown.

AR689943 873 bp DNA linear PAT 12-SEP-2005  
Sequence 13 from patent US 6905867.

ORGANISM Unknown.  
Unclassified.  
1 (bases 1 to 873)  
REFERENCE Gilbert,M. and Wakarchuk,W.W.  
AUTHORS Nucleic acids encoding polypeptides with .beta.1,3-galactosyl  
TITLE transferase activity  
JOURNAL Patent: US 6905867-A 13 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source Location/Qualifiers  
1..873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 99.3%; Score 866.6; DB 2; Length 873;  
Best Local Similarity 99.5%; Pred. No. 1e-243;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGTGATGTTTAGATGTAATCAATTTTATTGGAATTAATCTAT 120  
DB 61 CTACCAATGATTTTGTGATGTTTAGATGTAATCAATTTTATTGGAATTAATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTCTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTCTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAATTAATGTTGTTCTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAATTAATGTTGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACAACTTTTAAAGAAATTGATTATTCAGG 300  
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACAACTTTTAAAGAAATTGATTATTCAGG 300  
QY 301 GATGCTCATTTGGGATGATGTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 360  
DB 301 GATGCTCATTTGGGATGATGTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 360  
QY 361 TTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTATTAAGAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTATTAAGAAATGGGTCA 480  
QY 481 TCTTATGCTTTGGATATGATTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 540  
DB 481 TCTTATGCTTTGGATATGATTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 540  
QY 541 GATCGCTCAGTATATCGGACATAGTAAATAACAGATATAAGAAATTTAGAAATTTCTA 600  
DB 541 GATCGCTCAGTATATCGGACATAGTAAATAACAGATATAAGAAATTTAGAAATTTCTA 600  
QY 601 GAAAAAATTACAAAAATAAACTATATTTGCTTATGCTTAACTAAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTACAAAAATAAACTATATTTGCTTATGCTTAACTAAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATCATACAGAAATAAATTAACACTAC 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATCATACAGAAATAAATTAACACTAC 720  
QY 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTCAAAAAATTT 780  
DB 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTCAAAAAATTT 780  
QY 781 AAAAAATTAATAAGAAATTTTATTTACAGTTGATAAAGATCTATTAGATTA 840  
DB 781 AAAAAATTAATAAGAAATTTTATTTTACAGTTGATAAAGATCTATTAGATTA 840

Db 781 AAAAAATTAATAAGAAATTTTATTTACAGTTGATAAAGATCTATTAGATTA 840  
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 14  
AR691839  
LOCUS AR691839 873 bp DNA linear PAT 13-SEP-2005  
DEFINITION Sequence 13 from patent US 6911337.  
ACCESSION AR691839  
VERSION AR691839.1 GI:74479550  
KEYWORDS Unknown.  
SOURCE Unclassified.  
ORGANISM 1 (bases 1 to 873)  
REFERENCE Gilbert,M. and Wakarchuk,W.W.  
AUTHORS Nucleic acids encoding .beta.-1,4-GalNac transferase  
TITLE Patent: US 6911337-A 13 28-JUN-2005;  
JOURNAL National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source Location/Qualifiers  
1..873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 99.3%; Score 866.6; DB 2; Length 873;  
Best Local Similarity 99.5%; Pred. No. 1e-243;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTGATTATTCAGG 60  
DB 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTGATTATTCAGG 60  
QY 61 CTACCAATGATTTTGTGATGTTTAGATGTAATCAATTTTATTGGAATTAATCTAT 120  
DB 61 CTACCAATGATTTTGTGATGTTTAGATGTAATCAATTTTATTGGAATTAATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTCTTGAGCAATACTAC 180  
DB 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTCTTGAGCAATACTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAATTAATGTTGTTCTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGCCGAATTAATGTTGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACAACTTTTAAAGAAATTGATTATTCAGG 300  
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACAACTTTTAAAGAAATTGATTATTCAGG 300  
QY 301 GATGCTCATTTGGGATGATGTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 360  
DB 301 GATGCTCATTTGGGATGATGTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 360  
QY 361 TTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
DB 361 TTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 420  
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTATTAAGAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTATTAAGAAATGGGTCA 480  
QY 481 TCTTATGCTTTGGATATGATTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 540  
DB 481 TCTTATGCTTTGGATATGATTTTAAACAACTTTTAAAGAAATTGATTATTCAGG 540  
QY 541 GATCGCTCAGTATATCGGACATAGTAAATAACAGATATAAGAAATTTAGAAATTTCTA 600  
DB 541 GATCGCTCAGTATATCGGACATAGTAAATAACAGATATAAGAAATTTAGAAATTTCTA 600  
QY 601 GAAAAAATTACAAAAATAAACTATATTTGCTTATGCTTAACTAAGTCTTTTACCAATTTT 660  
DB 601 GAAAAAATTACAAAAATAAACTATATTTGCTTATGCTTAACTAAGTCTTTTACCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATCATACAGAAATAAATTAACACTAC 720  
DB 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATCATACAGAAATAAATTAACACTAC 720  
QY 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTCAAAAAATTT 780  
DB 721 AAAGATATCTATACCTCTTAGTGAGGCTTTATGGAATTTTCAAAAAATTTTCAAAAAATTT 780  
QY 781 AAAAAATTAATAAGAAATTTTATTTACAGTTGATAAAGATCTATTAGATTA 840  
DB 781 AAAAAATTAATAAGAAATTTTATTTTACAGTTGATAAAGATCTATTAGATTA 840



GenCore version 5.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: July 19, 2006, 14:53:32 ; Search time 652.919 Seconds  
 (without alignments)  
 9322.405 Million cell updates/sec  
 Title: US-10-734-719-11  
 Perfect score: 873  
 Sequence: 1 atgaaaaagttattatgc.....agcattatttcaaggaaaa 873  
 Scoring table: IDENTITY NUC  
 Gapop 10.0 , Gapext 1.0  
 Searched: 5244920 seqs, 3486124231 residues  
 Total number of hits satisfying chosen parameters: 10489840  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : N\_Geneseq\_8.\*  
 1: Geneseqn1980s.\*  
 2: Geneseqn1990s.\*  
 3: Geneseqn2000s.\*  
 4: Geneseqn2001as.\*  
 5: Geneseqn2001bs.\*  
 6: Geneseqn2002as.\*  
 7: Geneseqn2002bs.\*  
 8: Geneseqn2003as.\*  
 9: Geneseqn2003bs.\*  
 10: Geneseqn2003cs.\*  
 11: Geneseqn2003ds.\*  
 12: Geneseqn2004as.\*  
 13: Geneseqn2004bs.\*  
 14: Geneseqn2005s.\*  
 15: Geneseqn2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	873	100.0	873	8	ABT13670 C. jejuni
2	866.6	99.3	873	8	ABT13671 C. jejuni
3	866.6	99.3	876	3	AA53726 Campylob
4	866.6	99.3	876	8	ABT13669 C. jejuni
5	847.4	97.1	876	3	AA53724 Campylob
6	847.4	97.1	876	8	ABT13667 C. jejuni
7	845.8	96.9	876	13	ADU77566 Sialytran
8	845.8	96.9	12219	13	ADU77595 Campylob
9	836.2	95.8	11474	3	AA53720 LPS core
10	836.2	95.8	11474	8	ABT13665 Campylob
11	834.6	95.6	876	3	AA53721 Campylob
12	834.6	95.6	876	8	ABT13666 C. jejuni
13	820.2	94.0	876	3	AA53725 Campylob
14	820.2	94.0	876	8	ABT13668 C. jejuni
15	314.8	36.1	1293	2	AA225693
C 16	285.2	32.7	110000	2	AA242063_03
C 17	281	32.2	191996	13	ADT05647
C 18	264	30.2	13379	13	ADT05530

C 19	264	30.2	117576	13	ADT05736
C 20	264	30.2	349980	13	ADT05648
C 21	260.4	29.8	4277	13	ADT05421
C 22	259.8	29.8	909	8	ACC71694
C 23	88.6	10.1	2482	14	ADZ71091
C 24	87.6	10.0	8056	8	ABZ10246
C 25	86.2	9.9	8056	8	ABZ10246
C 26	86.2	9.9	19087	6	ABL32793
C 27	80.6	9.2	15548	6	ABL34155
C 28	80.4	9.2	8056	8	ABZ10100
C 29	78	8.9	8056	8	ABZ10100
C 30	77.8	8.9	158001	12	ADL17884
C 31	75.6	8.7	18988	4	AAS46342
C 32	75.6	8.7	18988	6	ABL32701
C 33	75.6	8.7	18988	6	ABL34509
C 34	75.6	8.7	18988	6	ABL70204
C 35	75.6	8.7	18988	7	ADS99770
C 36	74.8	8.6	1554	14	ADZ70999
C 37	73.4	8.4	6465	6	ABL32985
C 38	72.4	8.3	1501	8	ABZ10188
C 39	72.4	8.3	1501	10	ADB84162
C 40	72.4	8.3	5979	4	AAS45313
C 41	72.4	8.3	5979	6	ABK28152
C 42	72	8.2	17131	6	ABL33053
C 43	71.8	8.2	6652	6	ABQ66987
C 44	71.8	8.2	12592	6	AAS61102
C 45	71.8	8.2	17967	6	ABL33014

## ALIGNMENTS

RESULT 1  
 ABT13670  
 ID ABT13670 standard; DNA; 873 BP.

XX AC ABT13670;

XX DT 07-FEB-2003 (first entry)

XX DE C. jejuni bifunctional sialtransferase cstII coding sequence #5.

XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.

XX PN WO200274942-A2.

XX PD 26-SEP-2002.

XX XX 22-FEB-2002; 2002WO-CA000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX PA (CANA ) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX DR WPI; 2003-040554/03.

XX PT P-PSDB; ABJ18484.

XX PS New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.

XX PS Disclosure; Page 98-99; 107pp; English.

XX CC The invention comprises the amino acid and coding sequences of

CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; Galnac (N-  
 CC Acetylglactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthetase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention  
 XX  
 XX  
 SQ Sequence 873 BP; 350 A; 118 C; 110 G; 295 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 873; DB 8; Length 873;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-133; Indels 0; Gaps 0;  
 Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
 DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
 QY 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
 DB 61 CTACCAATGATTTTGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120  
 QY 121 CTTGGTAAATGCAAGACGCTGTTTACACCCCTGGTCTCTTTTGGCAATACACTAC 180  
 DB 121 CTTGGTAAATGCAAGACGCTGTTTACACCCCTGGTCTCTTTTGGCAATACACTAC 180  
 QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAATAATATGTTCTTAAT 240  
 DB 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAATAATATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCC 300  
 DB 241 TACAACCAAGCTATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATATTTTCC 300  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAATTTTAA 360  
 DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAAATTTTAA 360  
 QY 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCAGTAGCC 420  
 DB 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCAGTAGCC 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
 QY 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 DB 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 QY 541 GATGCTCACAATATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
 DB 541 GATGCTCACAATATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA 600  
 QY 601 GAAAAAATCTTACAAAATAAAATATAATGCTTATGCTTCTTAACAGTCTTTTACGAAATTT 660  
 DB 601 GAAAAAATCTTACAAAATAAAATATAATGCTTATGCTTCTTAACAGTCTTTTACGAAATTT 660  
 QY 661 ATAGAATCTAGCGCAATTTAAATTTTAAATTTTATCATACAGAAAAAATACTACACT 720  
 DB 661 ATAGAATCTAGCGCAATTTAAATTTTAAATTTTATCATACAGAAAAAATACTACACT 720  
 QY 721 AAGATATCTATCATACCTTCTAGTGAGCTTATCGGAAATTTTCAAAAATATTAATTTT 780  
 DB 721 AAGATATCTATCATACCTTCTAGTGAGCTTATCGGAAATTTTCAAAAATATTAATTTT 780  
 QY 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840

DB 781 AAAAAAATAAAATTTAAAGAAATGTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
 QY 841 CCTAGTGATATAAAGCAATTTATTTCAAAGGAAAA 873  
 DB 841 CCTAGTGATATAAAGCAATTTATTTCAAAGGAAAA 873  
 RESULT 2  
 ABT13671  
 ID ABT13671 standard; DNA; 873 BP.  
 AC ABT13671;  
 XX  
 XX 07-FEB-2003 (first entry)  
 XX C. jejuni bifunctional sialtransferase cstII coding sequence #6.  
 XX Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW Galnac transferase; N-Acetylglactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 XX Campylobacter jejuni.  
 XX WO200274942-A2.  
 XX 26-SEP-2002.  
 XX 22-FEB-2002; 2002WO-CA000229.  
 XX 21-MAR-2001; 2001US-00816028.  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 XX Gilbert M, Wakarchuk WW;  
 FI WPI; 2003-040554/03.  
 DR P-PSDB; ABJ18485.  
 XX  
 XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Disclosure; Page 99; 107pp; English.  
 XX  
 XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; Galnac (N-  
 CC Acetylglactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention  
 XX  
 SQ Sequence 873 BP; 353 A; 117 C; 107 G; 296 T; 0 U; 0 Other;  
 Query Match 99.3%; Score 866.6; DB 8; Length 873;  
 Best Local Similarity 99.5%; Pred. No. 1.2e-132;  
 Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
 DB 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60





181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAATAATATATGTTCTAAT 240  
 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCT 300  
 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360  
 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360  
 361 TTTTCAGGAATTTTCAATCAAGAATTAACCTCAGGGCTATATGTCAGTAGCC 420  
 361 TTTTCAGGAATTTTCAATCAAGAATTAACCTCAGGGCTATATGTCAGTAGCC 420  
 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 541 GATGCTCACATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCTA 600  
 541 GATGCTCGCATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCTA 600  
 601 GAAAAAATCTTACAAATAAACTATATGCTTATGTCCTTAAACAGTCTTTTACCAATTT 660  
 601 GAAAAAATCTTACAAATAAACTATATGCTTATGTCCTTAAACAGTCTTTTACCAATTT 660  
 661 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAAAATACTACAT 720  
 661 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAAAATACTACAT 720  
 721 AAAGATATACATACCTTCTAGTACGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
 721 AAAGATATACATACCTTCTAGTACGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
 781 AAAAAATTAATAATTAAGAAATCTTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
 781 AAAAAATTAATAATTAAGAAATCTTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840  
 841 CCTAGTATATAAGCATTTTCAAGGAAAA 873  
 841 CCTAGTATATAAGCATTTTCAAGGAAAA 873

RESULT 4

ABT13669  
 ID ABT13669 standard; DNA; 876 BP.  
 XX  
 AC ABT13669;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE C. jejuni bifunctional sialtransferase cstII coding sequence #4.  
 XX  
 KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 XX  
 PN WO200274942-A2.  
 XX  
 XX 26-SEP-2002.  
 XX  
 PF 22-FEB-2002; 2002WO-CA000229.  
 XX

21-MAR-2001; 2001US-00816028.  
 (CAN ) NAT RES COUNCIL CANADA.  
 Gilbert M, Wakarchuk WW;  
 WPI; 2003-040554/03.  
 P-PSDB; ABJ18482.  
 New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.  
 Disclosure; Page 98; 107pp; English.  
 The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an: acyltransferase; glycosyltransferase; GalNAc (N-Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present DNA sequence represents a Campylobacter jejuni gene of the invention  
 Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;  
 Query Match 99.3%; Score 866.6; DB 8; Length 876;  
 Best Local Similarity 99.5%; Pred. No. 1.2e-132;  
 Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTCTCGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
 DB 1 ATGAAAAAGTTATTATTCTCGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
 QY 61 CTACCAATGATTTTGTATGTTATGATGTTATCAATTTTATTTTGAAGATAATACTAT 120  
 DB 61 CTACCAATGATTTTGTATGTTATGATGTTATCAATTTTATTTGAAGATAATACTAT 120  
 QY 121 CTGTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCTTTTGGAGCAATACTAC 180  
 DB 121 CTGTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCTTTTGGAGCAATACTAC 180  
 QY 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGCCGAATAATATGTTCTTAAT 240  
 DB 181 ACTTTAAAAACATTTAATCCAAATCAAGAATATGAGCCGAATAATATGTTCTTAAT 240  
 QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCTCT 300  
 DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCTCT 300  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360  
 DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360  
 QY 361 TTTTCAGGAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420  
 DB 361 TTTTCAGGAATTTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCAGTAGCC 420  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
 DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGTCA 480  
 QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 DB 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 QY 541 GATGCTCACATATATCGGACATAGTAAATATACAGATATATAAGCTTTAGAAATTTCTA 600

```

Db 541 GATCGCTCGCACTATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 660
Db 601 GAAAAAATTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 720
QY 721 AAAGATATACATCATCTCTAGTCAGGCTTATGGAATTTTCAAAATTTTCAAAATTTT 780
Db 721 AAAGATATACATCATCTCTAGTCAGGCTTATGGAATTTTCAAAATTTTCAAAATTTT 780
QY 781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
Db 781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

```

RESULT 5

```

ID AAA53724 standard; DNA; 876 BP.
AC AAA53724;
XX
XX
DT 15-SEP-2003 (revised)
DT 22-DEC-2000 (first entry)
XX
XX Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.
XX
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside; ds.
XX
OS Campylobacter jejuni; O:10 serotype.
XX
FH Location/Qualifiers
FT 1..876
FT /*tag= a
FT /product= "alpha-2,3-sialyltransferase"
FT
XX WO200046379-A1.
XX
XX 10-AUG-2000.
XX
XX 01-FEB-2000; 2000WO-CA000086.
XX
XX 01-FEB-1999; 99US-0118213P.
XX 31-JAN-2000; 2000US-00495406.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
XX WPI; 2000-524418/47.
XX P-PSDB; AAY97210.
XX
XX Novel glycosyltransferase polypeptides and polynucleotides useful for
XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic
XX reagents and as immunogen for producing antibodies.
XX
XX Claim 6; Page 92; 120pp; English.
XX
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is
XX useful for synthesizing sialylated oligosaccharide such as ganglioside,
XX lysoganglioside or their mimics. Glycosyltransferases are useful for

```

```

CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
CC other oligosaccharides that have biological activity. The enzymes and
CC nucleic acids that encode them are useful for studies of the pathogenesis
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
CC jejuni and the nucleic acids are used as probes to study expression of
CC genes involved in ganglioside mimetic synthesis. Antibodies raised
CC against the glycosyltransferases are also useful for analyzing the
CC expression patterns of these genes involved in pathogenesis. The nucleic
CC acids are also useful for designing antisense oligonucleotides for
CC inhibiting expression of the Campylobacter enzymes that are involved in
CC the biosynthesis of ganglioside mimics that can mask the pathogens from
CC the host's immune system. The oligosaccharides are useful as diagnosing
CC reagents or as therapeutics and as immunogens for producing antibodies.
CC Bacterial glycosyltransferase can be used to catalyze the formation of
CC oligosaccharides that are identical to the corresponding mammalian
CC structures and are easier and less expensive to produce in large
CC quantity, compared to the mammalian glycosyltransferase. The bacterial
CC origin of the enzymes facilitates expression of large quantities of the
CC enzymes using relatively inexpensive prokaryotic expression systems.
CC (Updated on 15-SEP-2003 to standardise OS field)
XX
XX Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;
SQ

```

```

Query Match 97.1%; Score 847.4; DB 3; Length 876;
Best Local Similarity 98.2%; Pred. No. 1.7e-129;
Matches 857; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 ATGAAAAAGTTTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
Db 1 ATGAAAAAGTTTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
QY 61 CTACCAATGATTTTGTATGTTATGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTTGTATGTTATGATGTAATCAATTTTATTTGAAGATAAATCTAT 120
QY 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTGGTGTCTTCTTTGAGCAATCTAC 180
Db 121 CTTGGTAAAAATGCAAAACAGTGTTTTACACCCCTGGTGTCTTCTTTGAGCAATCTAC 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATGTGTCTTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATGTGTCTTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360
QY 361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Db 361 TTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGTATACCAAAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Db 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
QY 601 GAAAAAATTTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATTT 660
Db 601 GAAAAAATTTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 720

```

Db 661 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATACTACACT 720  
 Qy 721 AAAGATATACATACACCTCTAGAGGCTTATGCGAAATTTTCAAAAAATATTAAATTT 780  
 Db 721 AAAGATATACATACACCTCTAGAGGCTTATGCGAAATTTTCAAAAAATATTAAATTT 780  
 Qy 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
 Db 781 AAAAAATATAAAATTAAGAAAAATATTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
 Qy 841 CCTAGTGATATAAAGCATTTATTTCAAAAGGAAAA 873  
 Db 841 CCTAGTGATATAAAGCATTTATTTCAAAAGGAAAA 873

RESULT 6

ABT13667  
 ID ABT13667 standard; DNA; 876 BP.

AC ABT13667;

DT 07-FEB-2003 (first entry)

DE C. jejuni bifunctional sialtransferase cstII coding sequence #2.

KW Gene; ds; enzyme; Gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.

OS Campylobacter jejuni.

XX WO200274942-A2.

XX 26-SEP-2002.

XX 22-FEB-2002; 2002WO-CA000229.

XX 21-MAR-2001; 2001US-00816028.

XX (CANADA) NAT RES COUNCIL CANADA.

XX Gilbert M, Wakarchuk WW;

XX WPI; 2003-040554/03.

XX P-PSDB; ABJ18480.

XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.

XX Claim 8; Page 96-97; 107pp; English.

XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-  
 CC Acetylglucosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni  
 CC gene of the invention

XX Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

XX Query Match 97.1%; Score 847.4; DB 8; Length 876;

Best Local Similarity 98.2%; Pred. No. 1.7e-129;  
 Matches 857; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
 Qy 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGCTTTAAAAAGAAATGATTATTCAAGG 60  
 Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGCTTTAAAAAGAAATGATTATTCAAGG 60  
 Qy 61 CTACCAATGATTTTGTATGTTATGATGTTAATCAATTTTATTTTGAAGATAAATACTAT 120  
 Db 61 CTACCAATGATTTTGTATGTTATGATGTTAATCAATTTTATTTTGAAGATAAATACTAT 120  
 Qy 121 CTGTTAAAAATGCAAGCAGTGTGTTTACACCCCTGCTTCTTTGAGCAATACTACTAC 180  
 Db 121 CTGTTAAAAATTCAAAGCAGTATTTTACAAATCTCTGCTCTTTTGAACAATACTACTAC 180  
 Qy 181 ACTTTAAAAATTTAAATCCAAATCAAGAAATATGAGACCGAACTAATATATGTTCTAAT 240  
 Db 181 ACTTTAAAAATTTAAATCCAAATCAAGAAATATGAGACCGAACTAATATATGTTCTAAT 240  
 Qy 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCCT 300  
 Db 241 TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTCCT 300  
 Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTTAAATGCTTTATTATAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTTAAATGCTTTATTATAA 360  
 Qy 361 TTTCAAGAAATTTTATTTCAATCAAAAGAAATTTTACCTCAGGGTCTATATGTCAGTAGCC 420  
 Db 361 TTTCAAGAAATTTTATTTCAATCAAAAGAAATTTTACCTCAGGGTCTATATGTCAGTAGCT 420  
 Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAATGGGTCA 480  
 Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 Qy 541 GATGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600  
 Db 541 GATGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600  
 Qy 601 GAAAAAATTTACAAAAATAAACTATATTGCTTTATGCTTACCAAGCTTTTACCAATTTT 660  
 Db 601 GAAAAAATTTACAAAAATAAACTATATTGCTTTATGCTTACCAAGCTTTTACCAATTTT 660  
 Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720  
 Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720  
 Qy 721 AAAGATATACATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATAATTTT 780  
 Db 721 AAAGATATACATACCTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATAATTTT 780  
 Qy 781 AAAAAATATAAAATTTAAAGAAATGTTTATTACAAAGTTTCATAAAGATCTATTAAAGATTA 840  
 Db 781 AAAAAATATAAAATTTAAAGAAATGTTTATTACAAAGTTTCATAAAGATCTATTAAAGATTA 840  
 Qy 841 CCTAGTGATATAAAGCATTTATTTCAAAAGGAAAA 873  
 Db 841 CCTAGTGATATAAAGCATTTATTTCAAAAGGAAAA 873

RESULT 7

ADU77566

ID ADU77566 standard; DNA; 876 BP.

XX AC ADU77566;

XX DT 10-FEB-2005 (first entry)

XX DE Sialyltransferase (cstII).

XX gene amplification; campylobacter infection; Sialytransferase; estII; ds.  
 XX Campylobacter jejuni.  
 XX WO2004101817-A1.  
 XX 25-NOV-2004.  
 XX 19-MAY-2004; 2004WO-AU000654.  
 XX 19-MAY-2003; 2003AU-00902445.  
 XX (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
 XX Shi F, Fry BN, Coloe PJ;  
 XX WPI; 2004-834013/82.  
 XX Identifying Campylobacter in a sample to comprises comparing the profile  
 PT of an amplification product to a comparator Campylobacter species after  
 PT performing nucleic acid amplification of the wla gene cluster.  
 XX Claim 6; SEQ ID NO 13; 75pp; English.  
 XX The invention describes a method of identifying Campylobacter in a sample  
 CC comprising comparing the profile of an amplification product to a  
 CC comparator Campylobacter species after performing nucleic acid  
 CC amplification. Identifying Campylobacter in a sample comprises: (a)  
 CC performing nucleic acid amplification by contacting the sample with a  
 CC pair of nucleic acid primers, where at least one nucleic acid primer is  
 CC derived from the wla gene cluster of Campylobacter, for a time and under  
 CC conditions for generation of an amplification product comprising the wla  
 CC gene cluster of Campylobacter or its portion; (b) preparing a profile of  
 CC the amplification product; and (c) comparing the profile to a profile of  
 CC a comparator Campylobacter species, strain or type, and thus, determining  
 CC the identity of Campylobacter in the sample. Also described are: (a) a  
 CC pair of nucleic acid primers, where at least one nucleic acid primer is  
 CC derived from the wla gene cluster of Campylobacter; or (b) a pair of  
 CC nucleic acid primers for amplification of the wla gene cluster of  
 CC Campylobacter or its portion; or (c) a probe capable of hybridizing to  
 CC the wla gene cluster of Campylobacter. The method and kit are useful for  
 CC identifying and detecting the presence of Campylobacter in a sample, or  
 CC for diagnosing infections associated with Campylobacter. This sequence  
 CC represents Sialytransferase, encoded by a gene in the Campylobacter  
 CC jejuni wla cluster used in the inventive method of detecting  
 CC Campylobacter jejuni infection.  
 XX  
 XX Sequence 876 BP; 356 A; 115 C; 104 G; 301 T; 0 U; 0 Other;  
 Query Match 96.9%; Score 845.8; DB 13; Length 876;  
 Best Local Similarity 98.1%; Pred. No. 3.1e-129;  
 Matches 856; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
 QY 1 ATGAAAAAGTTATTATTCGCGAAATGAGCAATTTTAAAGAAATTTAATGCTTATTTTAA 60  
 DB |||||||  
 QY 1 ATGAAAAAGTTATTATTCGCGAAATGAGCAATTTTAAAGAAATTTAATGCTTATTTTAA 60  
 DB |||||||  
 QY 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 DB |||||||  
 QY 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 DB |||||||  
 QY 121 CTTCGTAATAAATGCAAGCAGTGTGTTTACACCCCTGCTTCTCTTTGAGCAATACCTAC 180  
 DB |||||||  
 QY 121 CTTCGTAATAAATGCAAGCAGTGTGTTTACCAATCTGCTGCTTTTGTGAAACAATACTAC 180  
 DB |||||||  
 QY 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGAGACCGCACTAATTTATGTTCTTAAT 240  
 DB |||||||  
 QY 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGAGACCGCACTAATTTATGTTCTTAAT 240  
 DB |||||||  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTTCAGATATTTTCCCT 300  
 DB |||||||  
 QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTTCAGATATTTTCCCT 300  
 DB |||||||

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
 DB |||||||  
 QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA 360  
 DB |||||||  
 QY 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGACATAGCC 420  
 DB |||||||  
 QY 361 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGACATAGCC 420  
 DB |||||||  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATTTGATTTTATCAAAATCGGTC 480  
 DB |||||||  
 QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGAAATTTGATTTTATCAAAATCGGTC 480  
 DB |||||||  
 QY 481 TCTTATGCTTTTCATATACCAAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 DB |||||||  
 QY 481 TCTTATGCTTTTCATATACCAAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
 DB |||||||  
 QY 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
 DB |||||||  
 QY 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600  
 DB |||||||  
 QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTACAGCTCTTTTACGCAATTTT 660  
 DB |||||||  
 QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTTACAGCTCTTTTACGCAATTTT 660  
 DB |||||||  
 QY 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
 DB |||||||  
 QY 661 ATAGAACTAGCGCAAAATTTTAAATTTTAAATTTTATCATACAAGAAAAATAACTACACT 720  
 DB |||||||  
 QY 721 AAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780  
 DB |||||||  
 QY 721 AAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780  
 DB |||||||  
 QY 781 AAAAAATAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTA 840  
 DB |||||||  
 QY 781 AAAAAATAAAATTTAAAGAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTA 840  
 DB |||||||  
 QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
 DB |||||||  
 QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
 DB |||||||  
 RESULT 8  
 ADU77595  
 ID ADU77595 standard; DNA; 12219 BP.  
 XX AC ADU77595;  
 XX DT 10-FEB-2005 (first entry)  
 XX DE Campylobacter jejuni strain ATCC 43469 wla cluster.  
 XX KW gene amplification; campylobacter infection; wla; ds.  
 XX OS Campylobacter jejuni.  
 XX PN WO2004101817-A1.  
 XX PD 25-NOV-2004.  
 XX PF 19-MAY-2004; 2004WO-AU000654.  
 XX PR 19-MAY-2003; 2003AU-00902445.  
 XX PA (MICR-) MICROTECHNOLOGY CENT MANAGEMENT LTD.  
 XX Shi F, Fry BN, Coloe PJ;  
 XX WPI; 2004-834013/82.  
 XX Identifying Campylobacter in a sample to comprises comparing the profile  
 PT of an amplification product to a comparator Campylobacter species after  
 PT performing nucleic acid amplification of the wla gene cluster.

XX Disclosure; Fig 6; 75pp; English.

XX The invention describes a method of identifying Campylobacter in a sample

CC comprising comparing the profile of an amplification product to a

CC comparator Campylobacter species after performing nucleic acid

CC amplification. Identifying Campylobacter in a sample comprises: (a)

CC performing nucleic acid amplification by contacting the sample with a

CC pair of nucleic acid primers, where at least one nucleic acid primer is

CC derived from the wla gene cluster of Campylobacter, for a time and under

CC conditions for generation of an amplification product comprising the wla

CC gene cluster of Campylobacter or its portion; (b) preparing a profile of

CC the amplification product; and (c) comparing the profile to a profile of

CC a comparator Campylobacter species, strain or type, and thus, determining

CC the identity of Campylobacter in the sample. Also described are: (a) a

CC pair of nucleic acid primers, where at least one nucleic acid primer is

CC derived from the wla gene cluster of Campylobacter; or (b) a pair of

CC nucleic acid primers for amplification of the wla gene cluster of

CC Campylobacter or its portion; or (c) a probe capable of hybridizing to

CC the wla gene cluster of Campylobacter. The method and kit are useful for

CC identifying and detecting the presence of Campylobacter in a sample, or

CC for diagnosing infections associated with Campylobacter. This sequence

CC represents a Campylobacter jejuni strain ATCC 43469 wla cluster.

XX

SQ Sequence 12219 BP; 4778 A; 1598 C; 1745 G; 4098 T; 0 U; 0 Other; 1

Query Match 96.9%; Score 845.8; DB 13; Length 12219;

Best Local Similarity 98.1%; Pred. No. 3e-129;

Matches 856; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy	1	ATGAAAAAGTTATTTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG	60
Db	5995	ATGAAAAAGTTATTTCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGG	6054
Qy	61	CTACCAATGATTTTGTATTTAGATGTAATCAATTTATTTTGAAGATAACTAT	120
Db	6055	CTACCAATGATTTTGTATTTAGATGTAATCAATTTATTTTGAAGATAACTAT	6114
Qy	121	CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCTTTTGACCAATAC	180
Db	6115	CTTGGTAAAAATGCAAGCAGTGTTTTACCAATCTGGTCTTTTGTGACCAATAC	6174
Qy	181	ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGCACTTAATATGTTCTAAT	240
Db	6175	ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGCACTTAATATGTTCTAAT	6234
Qy	241	TACACCAAGCTCATCTAGAAATGCAAAATTTTCTGAAACTTTTACGATTATTTCT	300
Db	6235	TACACCAAGCTCATCTAGAAATGCAAAATTTTCTGAAACTTTTACGATTATTTCT	6294
Qy	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTTAA	360
Db	6295	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTTAA	6354
Qy	361	TTTCAGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATGTGTCAGTAGCC	420
Db	6355	TTTCAGAAATTTATCTCAATCAAGAAATACCTCAGGGGTCTATATGTGTCAGTAGCT	6414
Qy	421	ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA	480
Db	6415	ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA	6474
Qy	481	TCATTGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT	540
Db	6475	TCATTGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT	6534
Qy	541	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAGCTTTAGAAATTTCTA	600
Db	6535	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAGCTTTAGAAATTTCTA	6594
Qy	601	GAATAAATTTACAAATATAATATATGCTTATGCTCCTAACAGCTCTTTAGCAAAATTT	660
Db	6595	GAATAAATTTACAAATATAATATATGCTTATGCTCCTAACAGCTCTTTAGCAAAATTT	6654

Qy	661	ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAATAACTACACT	720
Db	6655	ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAGAAAAATAACTACACT	6714
Qy	721	AAAGATATACATACCTCTAGTGAGGCTTATGGAATTTTCAAAAAATATATATTT	780
Db	6715	AAAGATATACATACCTCTAGTGAGGCTTATGGAATTTTCAAAAAATATATATTT	6774
Qy	781	AAAAAAATAAATTTAAAGAAATCTTTTATCAAGTTGATAAAGATCTTTAAAGATTA	840
Db	6775	AAAAAAATAAATTTAAAGAAATCTTTTATCAAGTTGATAAAGATCTTTAAAGATTA	6834
Qy	841	CCTAGTGATATTAAGCAATTTATTTCAAAGGAAAA	873
Db	6835	CCTAGTGATATTAAGCAATTTATTTCAAAGGAAAA	6867
RESULT 9			
AAAS3720			
ID	AAAS3720	standard; DNA; 11474 BP.	
XX	AC	AAAS3720;	
XX	DT	15-SEP-2003 (revised)	
XX	DT	22-DEC-2000 (first entry)	
XX	DE	LPS core biosynthesis locus.	
XX	KW	Biosynthetic locus; biosynthesis; lipid A biosynthesis;	
KW	acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;		
KW	Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;		
KW	sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;		
KW	immunity; immunogen; ganglioside; ds.		
XX	Campylobacter jejuni; OH4384.		
OS			
XX	Key	Location/Qualifiers	
FT	CDS	350..1237	
FT		/tag= a	
FT		/product= "acyltransferase"	
FT		/note= "Open reading frame 2b"	
FT	CDS	1234..2490	
FT		/tag= b	
FT		/product= "glycosyl transferase"	
FT		/note= "Open reading frame 3a"	
FT	CDS	2786..3955	
FT		/tag= c	
FT		/product= "Glycosyl transferase"	
FT		/note= "Open reading frame 4b"	
FT	CDS	4025..5068	
FT		/tag= d	
FT		/product= "Beta 1,4-GalNAC transferase"	
FT		/note= "Open reading frame 5a"	
FT	CDS	5054..5959	
FT		/tag= e	
FT		/product= "Beta 1,2- galactosyltransferase"	
FT		/note= "Open reading frame 6a"	
FT	CDS	6048..6923	
FT		/tag= f	
FT		/product= "CstII sialyltransferase"	
FT		/note= "Open reading frame 7a"	
FT	CDS	6924..7964	
FT		/tag= g	
FT		/product= "sialic acid synthase"	
FT		/note= "Open reading frame 8a"	
FT	CDS	8020..9079	
FT		/tag= h	
FT		/product= "Sialic acid biosynthetic enzyme"	
FT		/note= "Open reading frame 9a"	
FT	CDS	9076..9741	
FT		/tag= i	
FT		/product= "CMP-sialic acid synthetase"	

FT CDS /note= "Open reading frame 10a"  
FT complement (10554..11366)  
FT /\*tag= j  
FT /product= "Lipooligosaccharide biosynthetic enzyme"  
FT /note= "Open reading frame 12a"  
XX WO200046379-A1.  
XX 10-AUG-2000.  
XX 01-FEB-2000; 2000WO-CA000086.  
XX 01-FEB-1999; 99US-0118213P.  
XX 31-JAN-2000; 2000US-00495406.  
XX (CAN ) NAT RES COUNCIL CANADA.  
XX Gilbert M, Wakarchuk WW;  
XX WPI: 2000-524418/47.  
XX P-PSDB; AAY97200, AAY97201, AAY97202, AAY97203, AAY97204, AAY97205,  
XX AAY97206.  
XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
XX reagents and as immunogen for producing antibodies.  
XX Claim 1; Page 86-90; 120pp; English.  
XX A reaction mixture for the synthesis of a sialylated oligosaccharide is  
XX useful for synthesising sialylated oligosaccharide such as ganglioside,  
XX lysoganglioside or their mimics. Glycosyltransferases are useful for  
XX chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
XX other oligosaccharides that have biological activity. The enzymes and  
XX nucleic acids that encode them are useful for studies of the pathogenesis  
XX mechanisms of organisms that synthesize ganglioside mimics, such as C.  
XX jejuni and the nucleic acids are used as probes to study expression of  
XX genes involved in ganglioside mimetic synthesis. Antibodies raised  
XX against the glycosyltransferases are also useful for analyzing the  
XX expression patterns of these genes involved in pathogenesis. The nucleic  
XX acids are also useful for designing antisense oligonucleotides for  
XX inhibiting expression of the Campylobacter enzymes that are involved in  
XX the biosynthesis of ganglioside mimics that can mask the pathogens from  
XX the host's immune system. The oligosaccharides are useful as diagnosing  
XX reagents or as therapeutics and as immunogens for producing antibodies.  
XX Bacterial glycosyltransferase can be used to catalyse the formation of  
XX oligosaccharides that are identical to the corresponding mammalian  
XX structures and are easier and less expensive to produce in large  
XX quantity, compared to the mammalian glycosyltransferase. The bacterial  
XX origin of the enzymes facilitates expression of large quantities of the  
XX enzymes using relatively inexpensive prokaryotic expression systems.  
XX (Updated on 15-SEP-2003 to standardise OS field)  
XX Sequence 11474 BP; 4506 A; 1495 C; 1613 G; 3860 T; 0 U; 0 Other;  
Query Match 95.8%; Score 836.2; DB 3; Length 11474;  
Best Local Similarity 97.4%; Pred. No. 1.1e-127;  
Matches 850; Conservative 0; Mismatches 23; Indels 0; Gaps 0;  
Qy 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6107  
Db 6048 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6107  
Qy 61 CTACCAATGATTTTGGATGATTATGATGTAATCAATTTTATTTGAGATAAATCTAT 120  
Db 6108 CTACCAATGATTTTGGATGATTATGATGTAATCAATTTTATTTGAGATAAATCTAT 6167  
Qy 121 CTTCGTAATAAATGCAAGCAGTGTATTACACCCCTGGTTCTTTCTTTGAGCAATCTAC 180  
Db 6168 CTTCGTAATAAATGCAAGCAGTGTATTACAACTCTATTCTTTTGTGACAACTACTAC 6227  
Qy 181 ACTTTAAACAATTTTAAATCCAAAATCAAGAAATGAGACCGCAACTAAATATGTTCTTAAT 240

Db 6228 ACTTTAAACAATTTAATCCAAATCAAGATATGAGACCGCAACTAATATATGTTCTTAAT 6287  
Qy 241 TCAACAACAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCT 300  
Db 6288 TCAACAACAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTTCCT 6347  
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACATTTAAAGAAATTTAATGCTTATTTAAA 360  
Db 6348 GATGCTCATTTGGGATATGATTTTAAACAACATTTAAAGAAATTTAATGCTTATTTAAA 6407  
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGCTGCAGTAGGCC 420  
Db 6408 TTTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGCTGCAGTAGGCC 6467  
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
Qy 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Db 6528 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6587  
Qy 541 GATCGCTCACATATATCGACATAGTAAATAACACAGATATATAAGCTTTAGAAATTTCTA 600  
Db 6588 GATAATTCACATATATCGACATAGTAAATAACACAGATATATAAGCTTTAGAAATTTCTA 6647  
Qy 601 GAAAAAACTTACAAAAATAAACTATATTCCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660  
Db 6648 GAAAAAACTTACAAAAATAAACTATATTCCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 6707  
Qy 661 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAATAACTACACT 720  
Db 6708 ATAGAACTAGCCCAAAATTTAAATTTTATCATACAGAAAAATAACTACACT 6767  
Qy 721 AAAGATATCTATCATACCTTCTAGTGAGGCTTATGAAATTTTCAAAAAATATTAATTTT 780  
Db 6768 AAAGATATCTATCATACCTTCTAGTGAGGCTTATGAAATTTTCAAAAAATATTAATTTT 6827  
Qy 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTTCAATAAAGATCTATTAAAGATTA 840  
Db 6828 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTTCAATAAAGATCTATTAAAGATTA 6887  
Qy 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
Db 6888 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 6920  
RESULT 10  
ABT13665  
ID ABT13665 standard; DNA; 11474 BP.  
XX  
XX AC ABT13665;  
XX  
XX DT 07-FEB-2003 (first entry)  
XX  
XX DE Campylobacter jejuni genomic lipooligosaccharide biosynthesis locus.  
XX  
XX KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GainAC transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate-sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
XX OS Campylobacter jejuni.  
XX  
XX FN WO200274942-A2.  
XX  
XX PD 26-SEP-2002.  
XX  
XX PF 22-FEB-2002; 2002WO-CA000229.  
XX  
XX PR 21-MAR-2001; 2001US-00816028.









```
Qy 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAACTAT 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAACTAT 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 CTTCGTAATAATGCAAGCAGTGTTTTACACCTCGTCTTCTTTTGGACATACTAC 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 CTTCGTAATAATGCAAGCAGTGTTTTACACCTCGTCTTCTTTTGGACATACTAC 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 ACTTTAAACATTTAATCCAAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 ACTTTAAACATTTAATCCAAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 241 TACACCAAGCTCATAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTCCT 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TACACCAAGCTCATAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTCCT 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTATAA 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 GATGCTCATTTGGGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTATAA 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 361 TTTCCAGAAATTTAATCCAAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 TTTCCAGAAATTTAATCCAAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAATTTCTA 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAATTTCTA 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 601 GAAAAAATTTTCAAAAATTAAGAAATTTGTTTATGCTTCTTACAGCTTTTAGCAAAATTT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 GAAAAAATTTTCAAAAATTAAGAAATTTGTTTATGCTTCTTACAGCTTTTAGCAAAATTT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAGAAAGAAATTAACACTACT 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAGAAAGAAATTAACACTACT 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 721 AAAGATATCTATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAATTTAATTTT 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 AAAGATATCTATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAATTTAATTTT 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 781 AAAAAATTAATAAAGAAATTTGTTTATPACAAAGTTGATAAAGATCTATTAGATTA 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 AAAAAATTAATAAAGAAATTTGTTTATPACAAAGTTGATAAAGATCTATTAGATTA 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 841 CCTAGTATTAAGCAATTTTCAAGGAAA 873
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 CCTAGTATTAAGCAATTTTCAAGGAAA 873
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 13

AAAS3725

ID AAAS3725 standard; DNA; 876 BP.

XX AC

XX AC

XX AC

DT 15-SEP-2003 (revised)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

XX 22-DEC-2000 (first entry)

Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
immunity; immunogen; ganglioside; ds.

Campylobacter jejuni; O:41 serotype.

Key Location/Qualifiers

CDS 1..876

/\*tag= a

/product= "alpha-2,3-sialyltransferase"

WO200046379-A1.

10-AUG-2000.

01-FEB-2000; 2000WO-CA000086.

01-FEB-1999; 99US-0118213P.

31-JAN-2000; 2000US-00495406.

(CANA ) NAT RES COUNCIL CANADA.

Gilbert M, Wakarchuk W;

WPI; 2000-524418/47.

P-PSDB; AAY97211.

Novel glycosyltransferase polypeptides and polynucleotides useful for  
biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
reagents and as immunogen for producing antibodies.

Claim 6; Page 94; 120pp; English.

A reaction mixture for the synthesis of a sialylated oligosaccharide is  
useful for synthesizing sialylated oligosaccharide such as ganglioside,  
lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
other oligosaccharides that have biological activity. The enzymes and  
nucleic acids that encode them are useful for studies of the pathogenesis  
mechanisms of organisms that synthesize ganglioside mimics, such as C.  
jejuni and the nucleic acids are used as probes to study expression of  
genes involved in ganglioside mimetic synthesis. Antibodies raised  
against the glycosyltransferases are also useful for analyzing the  
expression patterns of these genes involved in pathogenesis. The nucleic  
acids are also useful for designing antisense oligonucleotides for  
inhibiting expression of the Campylobacter enzymes that are involved in  
the biosynthesis of ganglioside mimics that can mask the pathogens from  
the host's immune system. The oligosaccharides are useful as diagnosing  
reagents or as therapeutics and as immunogens for producing antibodies.  
Bacterial glycosyltransferase can be used to catalyze the formation of  
oligosaccharides that are identical to the corresponding mammalian  
structures and are easier and less expensive to produce in large  
quantity, compared to the mammalian glycosyltransferase. The bacterial  
origin of the enzymes facilitates expression of large quantities of the  
enzymes using relatively inexpensive prokaryotic expression systems.  
(Updated on 15-SEP-2003 to standardise OS field)

Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Query Match 94.0%; Score 820.2; DB 3; Length 876;

Best Local Similarity 96.2%; Pred. No. 4.7e-125;

Matches 840; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60

Db 1 ATGAAAAAGTTATTATTCCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGA 60

Qy 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAACTAT 120

Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAACTAT 120

Qy	121	CTTGGTAAAAAATGC	AAAGCAGTGT	TTTTACACCCCTGGT	TTCTCTTTGAGCAATAC	TAAC	180
Db	121	CTTGGTAAAAAATGC	AAAGCAGTGT	TTTTACACCCCTGGT	TTCTCTTTGAGCAATAC	TAAC	180
Qy	181	ACTTTAAACATTTAA	TCAAAATCAAGA	TATGAGCCGAACT	TAATATGTTGTTCT	TAAT	240
Db	181	ACTTTAAACATTTAA	TCAAAATCAAGA	TATGAGCCGAACT	TAATATGTTGTTCT	TAAT	240
Qy	241	TACAACCAAGCTCAT	CTCTAGAAAAAT	GTGTAAAAACCT	TTTTTACGATTA	TTTTGCT	300
Db	241	TTTAAACCAAGCTCA	TCTAGAAAAAT	GTGTAAAAACCT	TTTTTACGATTA	TTTTGCT	300
Qy	301	GATGCTCATTTGGG	NATGATTTTTTTT	TAAAACAACCTTA	AGAAATTAATGCT	TATTTTAAA	360
Db	301	GATGCTCATTTGGG	NATGATTTTTTTT	TAAAACAACCTTA	AGAAATTAATGCT	TATTTTAAA	360
Qy	361	TTTTCACGAAATTT	ATTTTCAATCAAGA	ATTAACTCAGGGG	CTATATGTTGTCAGT	TAGCC	420
Db	361	TTTTCACGAAATTT	ATTTTCAATCAAGA	ATTAACTCAGGGG	CTATATGTTGTCAGT	TAGCC	420
Qy	421	ATAGCCCTTAGGAT	CAAAAGAAATTT	TATCTTTCGGAA	TTGATTTTTATCAAA	ATGGGTCA	480
Db	421	ATAGCCCTTAGGAT	CAAAAGAAATTT	TATCTTTCGGAA	TTGATTTTTATCAAA	ATGGGTCA	480
Qy	481	TCTTATGCTTTTGA	TACAACCAAGAA	ATCTTTTAAACCT	ATAGCCCTGAT	TTTTTAAAAAT	540
Db	481	TCTTATGCTTTTGA	TACAACCAAGAA	ATCTTTTAAACCT	ATAGCCCTGAT	TTTTTAAAAAT	540
Qy	541	GATCGCTCACACTA	TATCGGCATAGT	AAAAATACAGAT	ATAAAGCTTTAGA	ATTTCTTA	600
Db	541	GATCGCTCACACTA	TATCGGCATAGT	AAAAATACAGAT	ATAAAGCTTTAGA	ATTTCTTA	600
Qy	601	GAIAAAAACTTTACA	AAAAATAAAACT	TATATGCTTTAT	CTCTTAAACAGT	CTTTTACCAAA	660
Db	601	GAIAAAAACTTTACA	AAAAATAAAACT	TATATGCTTTAT	CTCTTAAACAGT	CTTTTACCAAA	660
Qy	661	ATAGAACTAGCGCC	CAAAATTTTAAAT	TTTATCATACA	AGAAAAAAATAA	CTACACT	720
Db	661	ATAGAACTAGCGCC	CAAAATTTTAAAT	TTTATCATACA	AGAAAAAAATAA	CTACACT	720
Qy	721	AAAGATATACCTCA	TACCTCTAGTGAGG	CTTATGGAATTTT	TGAAAAATTTTCA	AAAAATATTTA	780
Db	721	AAAGATATACCTCA	TACCTCTAGTGAGG	CTTATGGAATTTT	TGAAAAATTTTCA	AAAAATATTTA	780
Qy	781	AAAAAATAAAAAAT	TAAGAAAAATGTT	TATTAACAAGTT	GTATAAAGAGT	CTATTAAGATTA	840
Db	781	AAAAAATAAAAAAT	TAAGAAAAATGTT	TATTAACAAGTT	GTATAAAGAGT	CTATTAAGATTA	840
Qy	841	CTTAGTGATATAAA	AGCATATTTTCA	AAAGGAAAA	873		
Db	841	CTTAGTGATATAAA	AGCATATTTTCA	AAAGGAAAA	873		

RESIII.T 14

ABT13668

ABT13668  
ID ABT13668 standard; DNA: 876 BP.

XX

AC ABT

11

DT 07-FEB-2003 (first entry)

C. jejunii bifunctional sialtransferase cstII coding sequence #3.

XX	Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW	GalNAc transferase; N-Acetyl galactosamine transferase;
KW	galactosyltransferase; sialyltransferase; sialic acid synthase;
KW	cytidine 5'-monophosphate sialic acid synthetase;
KW	UMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW	ganglioside mimetics; inflammation; tumour metastasis.

XX  
OS *Campylobacter jejuni*.

XXX  
-----T-----

PN WO200274942-A2.

26-SEP-2002.

22-FEB-2002: 2002WO-CA000229.

21-MAR-2001: 2001US-00816028.

(CANA ) NAT RES COUNCIL CANADA.

Gilbert M, Wakarchuk WW;

WPI; 2003-040554/03.

P-PSDB; ABJ18481.

New glycosyltransferases from *Campylobacter*, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.

Claim 8; Page 97; 107pp; English.

The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an: acetyltransferase; galactosyltransferase; GalNAC (N-Acetyl-galactosamine) transferase; glycosyltransferase; sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present DNA sequence represents a Campylobacter jejuni gene of the invention

Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Query Match 94.0%: Score 820.2: DB 8: Lenath 876:

Best Local Similarity 96.2%; pred. No. 4.7e-125;

Matches 840; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTGCTGGAAATGGACCAAGTTTAAAGAAA TTGATTATTC AAG 60

47  
Db  
1 ATGAAAAAAGTTATATTGTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGA 60

QV 61 CTA CCAATGATTTTGTATTTTATGATGTAATCAATTTTATTTTGAAGATAAATPACTAT 12

61 CTACCAATGATTTTGGTGTATTTAGATGCAATCAATTTTAAATTTTGAAGATAAATCTAT 12

Qv 121 CTTGGTAAAAAATGCAAGCAGTGTTTACACCCCTGGTTTCCTCTTGAGCAATACTAC 18

57  
121  
121  
18

181 ACUTTAAACATTTAATCCAAATCAAGAAATGAGACCGA ACTAA TTA TGTGTTCTAAT 24

47  
181 ACTTTAAACATTATTCGAAATCAAGAATGAGACCGAAGTAATCAATGTTCTTAAT 24

[illegible]

db 241 TTTTACCAAGCTCATCTAGAAAATCAAAA TTTTGTAAAAA CTTTTTACGATTA TTTTTCCT 30

QV 301 GATGCTCAATTTGGGATAGATTATTTTAAACAACCTAAAGCAATTAAGCTTATTTTAAA 36

301 GATGCTCATTTGGGATATGATTTTTTCAACCACTTAAAGAAATCAATGCTTATTTTAAA 36

361 TTTTCA CGAAATTTATTTCAATCAAGAA TTA CCTCAGGGGTCTATATGTGTGCGATAGCC 42

587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

421 ATAGCCCTAGGATACACAGAAA TTTATCTTTGGGAAATGATTTTATATCAAAATGGGCTCA 48

db



Search completed: July 19, 2006, 15:26:28  
Job time : 655.252 secs

GenCore version 5.1.1.9  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: July 19, 2006, 14:59:51 ; Search time 4154.24 Seconds  
 (without alignments)  
 11751.261 Million cell updates/sec  
 Title: US-10-734-719-11  
 Perfect score: 873  
 Sequence: 1 atgaaaaagtattattgc.....agcattatttcaaggaaaa 873  
 Scoring table: IDENTITY NUC  
 Gapop 10.0 , Gapext 1.0  
 Searched: 48236798 seqs, 27959665780 residues  
 Total number of hits satisfying chosen parameters: 96473596  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : EST:  
 1: gb\_est1:\*  
 2: gb\_est3:\*  
 3: gb\_est4:\*  
 4: gb\_est5:\*  
 5: gb\_est6:\*  
 6: gb\_hic:\*  
 7: gb\_est2:\*  
 8: gb\_est7:\*  
 9: gb\_est8:\*  
 10: gb\_est9:\*  
 11: gb\_gss1:\*  
 12: gb\_gss2:\*  
 13: gb\_gss3:\*  
 14: gb\_gss4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93.4	10.7	1101	14	CNS00EVL
2	89	10.2	1179	1	AJ927794
3	88.8	10.2	1365	10	DT959930
4	88.4	10.1	1007	9	DN566957
5	87.2	10.0	1359	9	DN685273
6	86.2	9.9	1101	14	CNS0039G
7	85.8	9.8	1096	1	AJ927821
8	85.8	9.8	1354	12	CG744717
9	85.2	9.8	1227	14	AG330010
10	84.8	9.7	1542	14	AG386981
11	84	9.6	976	13	CW949642
12	83.4	9.6	1210	12	CG749728
13	83	9.5	1258	13	CL508886
14	82.8	9.5	1268	14	AG347098
15	82.8	9.5	1324	14	AG376784
16	82.6	9.5	930	7	BB898645
17	82.4	9.4	734	14	CNS010MP
18	82.4	9.4	1297	10	DV778934
19	82	9.4	1123	5	CD325190

C 20	81.8	9.4	906	14	AG555384
C 21	81.8	9.4	1101	14	CNS00BOI
C 22	81.6	9.3	759	14	CNS06QXV
C 23	81.6	9.3	1101	14	CNS00EVL
C 24	81.2	9.3	1489	10	DT957227
C 25	81	9.3	1186	13	CL649642
C 26	81	9.3	1242	12	CL068807
C 27	80.8	9.3	947	12	CL112501
C 28	80.8	9.3	1225	14	CNS0161D
C 29	80.6	9.2	963	10	DT797234
C 30	80.6	9.2	1118	10	DM600040
C 31	80.2	9.2	1354	12	CG757970
C 32	80	9.2	1277	12	CC253231
C 33	79.4	9.1	806	5	CJ449754
C 34	79.4	9.1	1083	1	AJ926215
C 35	79.4	9.1	1248	9	DN808090
C 36	79.2	9.1	1201	14	CNS0167M
C 37	79	9.0	945	2	BM358135
C 38	79	9.0	1101	14	CNS00EVL
C 39	79	9.0	1260	14	AG430117
C 40	78.8	9.0	1133	1	AJ926090
C 41	78.8	9.0	1355	14	AG346348
C 42	78.8	9.0	1638	14	AG347262
C 43	78.6	9.0	949	5	CK424488
C 44	78.4	9.0	1348	12	CG749499
C 45	78.4	9.0	1544	10	DV780516

ALIGNMENTS

RESULT 1  
 CNS00EVL  
 LOCUS  
 DEFINITION  
 Drosophila melanogaster genome survey sequence T7 end of BAC:  
 BAC29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION  
 AL069706  
 VERSION  
 AL069706.1 GI:4949849  
 KEYWORDS  
 Drosophila melanogaster (fruit fly)  
 SOURCE  
 Drosophila melanogaster  
 ORGANISM  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE  
 1 (bases 1 to 1101)  
 AUTHORS  
 Direct Submission  
 TITLE  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 JOURNAL  
 Bp 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 COMMENT  
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
 FEATURES  
 Location/Qualifiers  
 1..1101  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR29B23"  
 /clone\_lib="RPCI-98"

1 (bases 1 to 1179)

REFERENCE  
AUTHORS

Pain, A., Renaud, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W.,  
Kerhervou, A., Aslett, M., Bishop, R., Bouchier, C., Cochet, M.,  
Coulson, R.M.R., Cronin, A. de Villiers, E., Fraser, A., Fosker, N.,  
Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F.,  
Larke, N., Lord, A., Maser, P., McKellar, S., Mooney, P., Morton, F.,  
Nene, V., O'Neill, S., Price, C., Quail, M.A., Rabinowitsch, E.,  
Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T.,  
Squares, R., Squares, S., Tivey, A., Walker, A.R., Woodward, J.,

238	QY	AATTACCAACGAGCTCACTAGAAAATGAAAATTTTGTAAAAACCTTTTTCAGATTATTTT	297
332	Db	ANNNNNNNNNNNNNNANAAAAANANNTNNAAAAATATANANNAATAAAAAT	391
298	QY	CCTGATGCTCATTTGGGATATCATTTTAAAAACAACCTAAAAGAATTTAAATGCTTATTTT	357
392	Db	NATAAAAAATTATNTATTTTAAATATTTTNNATATATAAATAAAAAAATATATATTAATAAAT	451
358	QY	AAATTTTCAGAAATTTATTTCAATCAAGAATTTACTCAGGGGTCTATATGTGTCAGTGA	417
452	Db	AAAAAAAATTAATAAAAAATTTAATATATTTAAATNAAATAATATNTTTTAAAAATATAATATA	511
418	QY	GCATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGG	477
512	Db	AAATATATAAAATTAATAATAATNATTTATAAAAATTTAAATAATTAATAAATAATTAAT	571
478	QY	TCATCTTATGCTTTTGATACCAACCAAGAAAATCTTTTAAAACTAGCCCTGATTTTAAA	537
572	Db	TAAAAANTAAATAAATTAATAAAAAATTAATAATAATTTAATAAATAAAAAATAATNATA	631
538	QY	AATGATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAANTT	597
632	Db	ATATTTTAAATTAATAAATAAAAAAATATNANATAAAATATATATAATTTAAANAAAT	691
598	QY	CTAGAAAAAATTTACAAAAATAAACTATATATGCTTATGCTCTAACAGAGCTTTTAGCAAT	657
692	Db	ATATATTAATAAATAATAATATATATATTTTATATATAATTTAAAAATTTAAAAATA	751
658	QY	TTTATAGAACTAGCGCCAAATTTAAATTTAAATTTTATCATACAAGAAAAAATAACTAC	717
752	Db	TTTTTATTAATTAATTTTATAATAATATATATAAAAAATTAATAATTTATTAATATATA	811
718	QY	ACTAAGATATACTCATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATTTAAT	777
812	Db	AAAAAATTAATAAATAAANAATTTTAAAAAATAAATAATATATATATATAAAAAANAAT	871
778	QY	TTTTAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTTGATAAAAAAGATCTATTAGA	837
872	Db	TAATATATAATTAATAATAATTAATAATNTTTATAAATAATATTAANAATATATATTA	931
838	QY	TTACCTAGTGATATTAAGCATTTATTTCAAGGAA	871
932	Db	TTAAAAATTAATAAATAAATTTAATTTAATTTTAAAAATAA	965



cells of the sea urchin embryo  
Development 128 (131), 2615-2627 (2001)  
11493577  
Contact: Erica Sodergren  
Human Genome Sequencing Center  
Baylor College of Medicine  
One Baylor Plaza, Houston, TX 77030, USA  
Tel: 713-798-7676  
Fax: 713-798-6977  
Email: ericas@bcm.tmc.edu  
NCBI Trace Archive: 496925179  
Insert Length: 1750 Std Error: 0.25  
Plate: 143 row: H column: 12.  
Location/Qualifiers  
1. 1007  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="PMCSPr2-143H12"  
/tissue\_type="embryo"  
/cell\_type="primary mesenchyme cells"  
/lab\_host="E.coli;"  
/clone\_lib="Sea Urchin primary mesenchyme cell cDNA  
library"  
/note="vector: pSPOR1; Site 1: NotI; Site 2: MluI; oligo  
dt priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match	10.1%;	Score 88.4;	DB 9;	Length 1007;
Best Local Similarity	46.8%;	Pred. No. 4.4e-05;		
Matches 392;	Conservative 0;	Mismatches 431;	Indels 15;	Gaps 3
Qy	36	TTTAAAGAAATTCATTATTC	CAAGGCTACCAAAATGATTTTC	GATGTATTTAGATGTAATCA 95
Db	165	TTTAAATATTTT	TTTAAAAAAA	AAAAAATTTAAATTTAA
Qy	96	ATTTTATTTGAAGATAAAT	TACTATCTTGGTAA	AAAAATGCAAGCAGTGTTTACACCCC 155
Db	225	AAATTTTTTTTTT	TAAAAA	AAAAAATTTTTTTTTTTTAAAAA 284
Qy	156	TGCTTCTCTTCGACCAAT	TACTACACTTTAAACATTTAAT	CCAAATCAAGATATGA 215
Db	285	TTTTTTTTTTTTTTT	TAAAAA	AAAGGGTTTTTTTTTGTGTTAATAATTTTAAAAA 344
Qy	216	GACCGAACTAATATATG	TGTTCTAATACAA	CCAAAGCTCATCTAGAAAAATGAAATTTTGT 275
Db	345	AAACCTTTTTTTTTT	TTTTTTTATATA	AAAAAATAATTTTTTTTTTAAAAA 404
Qy	276	AAAACTTTTACGATTA	TTTCTCGTGTCA	TTTGGGATATGATTTTTTTTAAACAACT 335
Db	405	AAAAATTTTATATTT	TAAAAATTTTAAATTTTTTTTTT	TAAATAAAATTTTTTTTTTAAAAA 464
Qy	336	TAAAGAAATTTAAGCT	TATTTTAAATTTACGAAATTTAT	TTCATCAAGAAATTTACCTC 395
Db	465	CCAAAAAATAATA	TATTTTTTTTTTATATTTTTTTTTT	TTTTTTTTTAAAAATTTTTTTTTT 524
Qy	396	AGGGGTCTATATGTG	TCAGTAGCCATAGCCCTAGGATAC	AAAGAAATTTATCTTTCGG 455
Db	525	AA-----	TTTAAAAATTTATATATATTTAT	TATTTATATAAAAAAATTTTAAAAA 579
Qy	456	AATTGATTTTATCA	AAATGGGTCA	TCTTTTGATACCAAAACAAGAAATCTTTT 515
Db	580	TTTAAATTTTAAAA	AAAAAATTTTTTTTTTTTTTTTTT	TTTTTTTTTTTAAATATTTTTT 639
Qy	516	AAACTAGCCCTGAT	TTTTTAAATGATCGCTCA	CACATATACGACATAGTAAAAATAC 575
Db	640	TTTTTTATTTTTCT	TTTCTAAAAA-----	AAAAAATTTTTTAAATTTAGAAAAA 691
Qy	576	AGATATAAAGCTT	TAGAAATTTCTGTA	AAAAAATTTACAAATAAACTATATTCGTATG 635
Db	692	AAATTTTAAAAA	AAAAATATAAAAAA	ATACCTTTTAAA-----
Qy	636	TCCTAACAGTCT	TTTTTAGCAATTTTATAG	AACTAGCGCAAAATTTAAATTTCAATTTTAT 695



AL063921  
AL063921.1 GI:4941778  
GSS.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [segr@genoscope.cns.fr](mailto:segr@genoscope.cns.fr))  
Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 006 EVRY cedex - FRANCE (E-mail : [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr))  
Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Genoscope.  
Direct Submission  
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr)  
Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr))

- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
- Web : [sequif@genoscope.cns.fr](mailto:sequif@genoscope.cns.fr)

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Osoegawa and

melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RCI-98 and was constructed by partial

The library is named *RPL1-98* and was constructed by partial *EcoRI* digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain *y2*; *cn bw sp*, the same strain used for the BDGP's *p1* and *EST* libraries. A more detailed description of the library and how to order individual *BDG clones*, the entire library, or

and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
Location/Qualifiers

```
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
```

```
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TET3"
```

ch 9.9%; Score 86.2; DB 14; Length 1101;  
l Similarity 18.5%; Pred. No. 0.0001;  
130; Conservative 299; Mismatches 272; Indels 2; Gaps 1;

70 GATTITGATGATTTAGTAGTAATCAATTTTATTTTGAGATAAATACTATCTTGTTAA 129  
| | : : : : | : | : | : | : | : | : | : | : | :  
94 GMATATANWWWTTTTTTTTTAWAAAAGAATAATTTWAAWAAAAAATTTWWAAAAWA 453

[illegible][illegible]

74. TTTTYYYYYWTTWTHHTMYTHAAHTTTWTHWTHYANWHTWHTWTHWAWHWT 633  
0.0 TTGGATATGATTTTTTAAACAACCTTAAGAAATTTAATGCTTATTTTAAATTTACAGAA 369

34 TTTAAATTTTCATCAAGAAATTAACCTCAGGGGTCTATATGTGTGCAGTAGCCATAGCCCTTA 429

94 CTCTCHCCVUUUHHVTAAHTTTHHWYAHYUWUYWUYUCTACT--YHYNHHNHUHW 751

80 GGATACAAGAAAATTATCTTTTCGGGAATTGATTTTTTATCAAATGGGCATCTTATGCT 489

002 AYHTWYAWABAMWMMHHAHYAAAGAAWAAATHTHHYHHTHTHYMMHHTYMYHYMYTCCYMC 811

```

Db      812 TYHCWHYYHTAYTCWTWTHHHMMWTWTHWYHHHTHHHTTTHWAWHHTWCHWWHATWT 871
Qy      550 CACTATACGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTAGAAAAA 609
Db      872 WATHCWACMTWHHWHHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMHMH 931
Qy      610 TACAAAATAAACAATATATGCTTATGCTTAACCTTAACAGTCTTTTACGAAAATTTATAGA 669
Db      932 HMEHHWHHWWATWMTWMTTMMCMCMHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCH 991
Qy      670 GCGCCAAATTTAAATTCACAAATTTTATCATACAGAAAAAATACTACACTAAAGATATA 729
Db      992 YMCCTCYCHCTWHYWTAYWAWTAHAMTTATWMMHWHAWHATWMMHWHAWHATWMMH 1051
Qy      730 CTCACTCTTCTAGTGGCTTATGGAATAATTTTCAAAAAATA 772
Db      1052 CHWTWVHTCTWYHHHTYHMMWAWMMWAWHWHHMYAHYHWAHH 1094

RESULT 7
AJ927821
LOCUS      1096 bp mRNA linear EST 10-JUN-2005
DEFINITION Theileria annulata merozoite Theileria annulata cDNA clone
ACCESSION tam027c04_glk, mRNA sequence.
VERSION    AJ927821
KEYWORDS   Theileria annulata
SOURCE     Theileria annulata
ORGANISM   Theileria annulata
REFERENCE  1 (bases 1 to 1096)
AUTHORS   Pain,A., Renauld,H., Berriman,M., Murphy,L., Yeats,C.A., Weir,W.,
Kerhornou,A., Aslett,M., Bishop,R., Bouchier,C., Cochet,M.,
Coulson,R.M.R., Cronin,A., de Villiers,E., Fraser,A., Fosker,N.,
Gardner,M., Goble,A., Griffiths-Jones,S., Harris,D.E., Katzer,F.,
Larke,N., Lord,A., Maser,P., McKellar,S., Mooney,P., Morton,P.,
Nene,V., O'Neill,S., Price,C., Quail,M.A., Rabinowitsch,E.,
Rawlings,N.D., Rutter,S., Saunders,D., Seeger,K., Shah,T.,
Squares,R., Squares,S., Tivey,A., Walker,A.R., Woodward,J.,
Dobelaere,D.A.E., Langsley,G., Rajandream,M.-A., McKeever,D.,
Shields,B., Tait,A., Barrall,B. and Hall,N.
TITLE      The genome of the host-cell transforming parasite Theileria
JOURNAL    annulata and a comparison with T. parva
COMMENT    Unpublished (2005)
Contact: Pain A
The Pathogen Sequencing Unit
The Wellcome Trust Sanger Institute
Genome Campus, CB10 1SA, UNITED KINGDOM
Merozoite cDNA library: Frank Katzer and Brian Shields, Division of
Veterinary Infection and Immunity, ICM, University of Glasgow, UK.
Location/Qualifiers
1..1096
/organism="Theileria annulata"
/mol_type="mRNA"
/isolate="Ankara (clone D7)"
/db_xref="taxon:5874"
/clone="tam027c04_glk"
/dev_stage="merozoite"
/lab_host="Bos taurus (cow)"
/clone_lib="Theileria annulata merozoite"
/notes="country: Turkey;Ankara"

ORIGIN
Query Match 9.8%; Score 85.8; DB 1; Length 1096;
Best Local Similarity 36.7%; Pred. No. 0.00012;
Matches 262; Conservative 0; Mismatches 446; Indels 6; Gaps 1;
Qy      160 TTCTTCTTTCAGCAATACATACACTTTAAACAACTTTAATCCAAATCAAGATATGAGACC 219
Db      307 TTTTNNNTGGANNNNNATTANNNNANNAAGTATNNNNNNNNAAGTAAATTAATAAAAAAN 366
Qy      220 GAACTAATATGTGTTCTTAATTACAACCAAGCTCATCTGAAAAATGAAAAATTTTGTAAAA 279

```

```

Db      367 ANANNAANGGTAAATGTTTAAATATAANNANNGNNNTTTATTTTAAATAAATANNAATANTG 426
Qy      280 ACTTTTTCAGGATTAATTTTCTGATGCTCATTTGGGATATGATTTTCTTTTAAACAACATTTAAA 339
Db      427 NNTNATATANATNTAATATATATAANATNNNTNTATTTTATATANAATAAANNTNNN 486
Qy      340 GAAATTAATGCTTAATTTTAAATTTTCAAGAAATTTATTTCAATCAAAAGAAATTCCTCAGGG 399
Db      487 NNANNAAGAGTAAATTAATAAANANNNNTGNTTAAANATTTAAATAATNANAAANNT 546
Qy      400 GTCTATATGTGTCAGTAGCCATAGCCCTAGAGATACAAAGAAATTTATCTTTTCGGGAATT 459
Db      547 NNAAAAATAAATTAATANNNAANAAAAANNNNAANAAAAANAAAAANNTTTTAAAT 606
Qy      460 GATTTTATCAAAATGGGTCATCTTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAAA 519
Db      607 TAAANTATATAATANNAAAATTTAANTAANNAAAAAANNAANTTANAAATTTATTAANAA 666
Qy      520 CTAGCCCTGATTTTAAAAATGATCGCTCACATATATCGGACATAGTAAAAATACAGAT 579
Db      667 TTANAAAA-----ANAAAAANNNAAAAAATTTTANANAANAATNNNATNNAAAAAN 720
Qy      580 ATAAAGCTTTAGAAATTTCTAGAAAAAACTTACAAAAATAAACTATATTTGCTTATGTCCT 639
Db      721 NATANNAANANNNNNNANNAATAAANNNNAATAAATAAATAAATAAATAAATAAATAAATAA 780
Qy      640 AACAGTCTTTAGCAAAATTTTATGAAGCTAGCGCCAAATTTTAAATTTCAATTTTATCATA 699
Db      781 AANNATATTTANAANNTTTTAAAAANNNANNAANNAANNAANNAANNAANNAANNAANNAANNA 840
Qy      700 CAAGAAAAATAACACACTAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATA 759
Db      841 AAAAAAANTANNTTNNNAATTTANNATNAAATNNAAAAATNAAAAAATAGTATTTCTANTAAA 900
Qy      760 TTTTCAAAAAATATTAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 819
Db      901 TAATTAATAAATNAAATAAAAAAGATATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
Qy      820 ATAAAGATCTTAAAGATTACCTAGTCATATAAGACATTAATTTCAAGGAAAA 873
Db      961 ANNAANATATANNAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1014

RESULT 8
CG744717/c
LOCUS      CG744717 1354 bp DNA linear GSS 24-OCT-2003
DEFINITION P037-2-F03.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION CG744717
VERSION    CG744717.1 GI:37965585
KEYWORDS   GSS.
SOURCE     Pristionchus pacificus
ORGANISM   Pristionchus pacificus
REFERENCE  1 (bases 1 to 1354)
AUTHORS   Rukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE      An integrated physical and genetic map of the nematode Pristionchus
JOURNAL    pacificus
PUBMED     Mol. Genet. Genomics 269 (5), 715-722 (2003)
COMMENT    12884007
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
1..1354
FEATURES
source

```

```

/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcorI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcorI and cloning into the BAC
vector."

ORIGIN
Query Match          9.8%; Score 85.8; DB 12; Length 1354;
Best Local Similarity 46.7%; Pred. No. 0.00012;
Matches 390; Conservative 0; Mismatches 437; Indels 8; Gaps 4;

QY 4 AAAAAAGTTATTATTGCTGGAAATGACCAAGCTTTAAAGAAATGATTATTCAAGGCTA 63
Db 1319 AAAAAAGTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1260
QY 64 CCAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTATCTT 123
Db 1259 AAAAAAATAATATATATATATATATATATATATATATATATATATATATATATAT 1200
QY 124 CGTAAAAATGCAAGAGCTGTTTACACCCCTGGTTCCTTTGAGCAATACTACACT 183
Db 1199 AATAAATAATATATATAAATAATATATTTTAAATATATTTTAAATAAATAAATAA 1140
QY 184 TTAACAACATTTAATCAAAATCAAGAAATATGACCGAAGCTAATATATGTTCTAATTAC 243
Db 1139 AATTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1084
QY 244 AACCAAGCTCATCTAGAAAAAGAAATTTTGTAAAACTTTTTCAGATTTTCTGAT 303
Db 1083 AATAAATTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1024
QY 304 GCTCATTTGGGATGATGTTTTTAAACAACCTTAAAGAAATTTAAAGCTTTTAAATTT 363
Db 1023 ATAAATATATATANTAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAA 965
QY 364 CACGAAATTTATTTCAATCAAGAATTAACCTCAGGGTCTATATGTCGAGTAGCCATA 423
Db 964 TATTATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 905
QY 424 GCCCTAGGATCAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCACT 483
Db 904 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 845
QY 484 TATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAATGAT 543
Db 844 AAAAAATATATAATAAATAAATTTATTTTNTTTTATATATAAATAAATAAATAAATAA 785
QY 544 CGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTAGAA 603
Db 784 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 725
QY 604 AAAAATTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 663
Db 724 AATATTATTCGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 665
QY 664 GAATAGCCGCAAAATTTAAAT-TCAAATTTATCATACAGAAAAATAAATAAATAAATAA 722
Db 664 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 605
QY 723 AGATATACCTCATCTCTAGTGGCTTATGGAATTTT-TCAAAAATAAATAAATAAATAA 780
Db 604 ATATTATTTTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 545
QY 781 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 835
Db 544 ANTHAATTTAATAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 490

```

```

LOCUS       AG430010          1227 bp    DNA    linear    GSS 21-DEC-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-304N17.TJ, genomic survey
sequence.
ACCESSION   AG430010
VERSION     AG430010.1  GI:48073073
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus (Japanese wild mouse)
ORGANISM   Mus musculus molossinus
REFERENCE   1
AUTHORS     Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
            Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriaki,K. and
            Shiroishi,T.
TITLE       Contribution of Asian mouse subspecies Mus musculus molossinus to
            genomic constitution of strain C57BL/6J, as defined by BAC-end
            sequence-SNP analysis
JOURNAL     Genome Res. 14 (12), 2439-2447 (2004)
PUBMED     15574823
REFERENCE   2 (bases 1 to 1227)
AUTHORS     Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE       Direct Submission
JOURNAL     Submitted (17-NOV-2003) Mashira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@c.riken.jp).
            Tsukuba Institute, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@c.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
R.Site 1 : pBACe3.6
R.Site 2 : EcorI.
Location/Qualifiers
1. 1227
    /organism="Mus musculus molossinus"
    /mol_type="genomic DNA"
    /sub_species="molossinus"
    /db_xref="taxon:57486"
    /clone="MSMg01-304N17.TJ"
    /sex="male"
    /tissue type="mixture of kidney and spleen"
    /clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match          9.8%; Score 85.2; DB 14; Length 1227;
Best Local Similarity 47.3%; Pred. No. 0.00015;
Matches 364; Conservative 0; Mismatches 383; Indels 22; Gaps 3;

QY 95 AATTTTATTTTGAAGATAAATACTATCTTGGTAAAAAATGCAAGCAGTGTTTTACACC 154
Db 1226 AATTTTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1167
QY 155 CTGGTTTCTCTTTCGAGCAATACTACACTTTTAAACATTTTAAATCAAGAAATCAAGAA 214
Db 1166 ATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1107
QY 215 AGACCGAATAAATATGTTCTCTTAATTAACAACCAAGCTCATCTAGAAAAATGAAA 274
Db 1106 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1047
QY 275 TAAAACTTTTACGATTTATTTTCTCGATGCTCATTTGGGATATGATTTTTTAAACA 334
Db 1046 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 987

```

```

Qy 335 TTAAAGATTTAAATCTTATTTAAATTTACGAAATTTATTTCAATCAAGAAATTACCT 394
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 986 AAAAAATAAATAATATATATAATTTTAAAAATAAATAAATAAATAAATAAATAAAT 927
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 395 CAGGGGTCTATATGTGTCAGTAGCCATAGCCCTAGGATACAAAGAAATTTATCTTCGG 454
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 926 AAT-----AAAATTTATAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 883
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 455 GAATGTATTTTATCAAAATGGGTATCTTATGCTTTTGTATACCAAAACAAGAAATCTTT 514
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 882 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 823
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 515 TAAACTAGCCCTGATTTTAAATAGTCTCCTACACTATATCGGACATAGTAAATA 574
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 822 ATAAATA---AATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 766
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 575 CAGATATAAAGCTTTTAAATTTCTAGAAAAAATCTTACAAAAATAAACTATATGCTTAT 634
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 765 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 706
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 635 GTCTAAACAGCTTTTATAGCAATTTTATAGAACTAGCCCAAAATTTAAATTTCAAAATTTA 694
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 705 ---ATTAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 649
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 695 TCATACAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 754
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 648 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 589
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 755 GAAATTTTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 814
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 588 AAAAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 529
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 815 AGTTGATAAAGATCTTATTAAGATTTACCTAGTGTATATAAAGCAATTTT 863
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 528 AATTAATAATATATATATATATATATATATATATATATATATATATATATATAT 480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AG386981/c
LOCUS
DEFINITION
Mus musculus molossinus DNA, clone:MSMg01-201G10.TJ, genomic survey
sequence.
ACCESSION
AG386981
VERSION
AG386981.1 GI:47998186
KEYWORDS
GSS.
SOURCE
Mus musculus molossinus (Japanese wild mouse)
ORGANISM
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Abe,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T.,
Ezawa,K., Saitou,N., Hattori,M., Sakaki,Y., Moriaki,K. and
Shiroishi,T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
15574823
2 (bases 1 to 1542)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suehiro-cho,Tsukumi-Ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center.
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199

COMMENT

```

e-mail: abe@rtc.riken.jp

#### PRIMERS

Sequencing : TJ

#### LIBRARY

Vector : DBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

#### FEATURES

##### source

Location/Qualifiers

1..1542

/organism="Mus musculus molossinus"

/mol\_type="genomic DNA"

/sub\_species="molossinus"

/db\_xref="taxon:57486"

/clone="MSMg01-201G10.TJ"

/sex="male"

/tissue type="mixture of kidney and spleen"

/clone\_lib="MSMg01 Mouse Male BAC Library"

#### ORIGIN

```

Query Match      9.7%; Score 84.8; DB 14; Length 1542;
Best Local Similarity 46.6%; Pred. No. 0.00017;
Matches 408; Conservative 0; Mismatches 453; Indels 14; Gaps 5;

Qy 5 AAAAAAGTTATTATTCCTGGAATGACCAAGCTTTAAAGAAATTTGATTTCAAGGCTAC 64
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1441 ATNAAATAAATANTANTATATATTAAGATAAATAAATAATATATATATATATAN 1382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 65 CAAATGATTTTGATCTATTAGATGTAATCAATTTTTCGAAGATAAATCTATCTTCTG 124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1381 AATAATTAATAATAAANNATAGTAATAATTAATAAATAAATAAATAAATAAATAA 1322
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 125 GTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCCTCTTTGAGCAATCTACACTT 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1321 ATTAATAATATATA-TATAATATAAATAATATATATTTTANAAATAAATAAATAA 1263
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 185 TAAACATTTTATCCAAATCAAGAATATGAGACCGAAGCTAATATATGTTCTTAATTACA 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1262 TAAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1206
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 245 ACCAGCTCATCTAGAAATGAAATTTTGTAAGAACTTTTACGATTTATTTTCCGTGATG 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1205 AATAATAATAATTTATATATATATATATTTTAAATAAATAAATAAATAAATAAATA 1146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 305 CTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTTAAATGCTTTTAAATTTTC 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1145 TTAAATAAATAATATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1086
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 365 ACGAAATTTTCAATCAAGAAATACCTCAGGGGCTATATG-TGTGACGTAGCCATA 423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1085 AATAATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 1026
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 424 GCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGTCTCT 483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1025 AAAATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 970
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 484 TATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAATGAT 543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 969 TATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAAT 910
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 CGCTCACACTATATCGGACATAGTAAATAATACAGATAAAGCTTTTAGAAATTTCTAGAA 603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 909 TATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 850
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 604 AAAACTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 849 AATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 790
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 664 GAACCTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAA- - - - -AAAAATAACTACA 718
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 789 AATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 730
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 719 CTAAAGATATATCTCATACCTTCTAGTAGGCTTTAGGAAAAATTTTCAAAAAATTTAAT 778
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db	305	AAATAAAATATAAAATATAAAATATAATAAAATAAATAAATAATATAA
Qy	444	TTATCTTCGGGAATGTATTTTATCAAAATCGGTCATCTTATGCTTTTGATACCAACA
Db	365	TTATAATAAAAAATAAATATGAAAAATATATTTTATCATATATAAAAAATAAAT
Qy	504	AGAAATCTTTTAAAACTAGCCCTTGATTTTAAAAATGATCGCTCACACTATATCGACA
Db	425	AAATTTTAAATTTTAAAAATTTATAAAATTTCTTAATTTAGATATAATATAATAAAATTT
Db	484	TTATCTTCGGGAATGTATTTTATCAAAATCGGTCATCTTATGCTTTTGATACCAACA

[illegible]

Accession	Sequence	Position
Db	ATATAAAAAAAAAATATAAATTTTAAATATATAATTAATATAATTTATTTAAATAATAA	604
Qy	TTCAAAATTTTATCATACAAGAAAAATAAATCTACTA----AGATATACTCATACCTT	739
Db	TTTATTTAAATTAATTTTATATATAATAATAATAATAATAATAATAATAATAATAAAT	684
Qy	CTAGTAGGCCTATGGAATAATTTTCCAAAAATATTAAATTTTAAAAAAATAAAAAATTTAAAG	799
Db	ATATTTATATTAATTTTATTTTAAATATAATATTTAAATTTTAAATTTTAAATTTTAAAT	724

[illegible]

DB	QY	860	ATTTCAAAGGAAAA	873
DB		785	AAATAAATTTAATA	798

RESULT 12

RES001.12	CG749728	1210 bp	DNA	linear	GSS 24-OCT-2000			
LOCUS	P044-1-C03.za	Ppa	ECORI	BAC Library	Pristionchus pacificus genomic			
DEFINITION	genomic survey sequence.							
ACCESSION	CG749728							
VERSION	CG749728.1	GI:37970654						
KEYWORDS	GSS.							
SOURCE	Pristionchus pacificus							
ORGANISM	Pristionchus pacificus							

**REFERENCE**

Buntjer, J., van der Meulen, M. and Sommer, R. J.  
Srinivasan, J., Sinz, W., Jeease, T., Wiggers-Perebolte, L., Jansen, K.,  
1 (bases 1 to 1210)  
Neodiplogasteridae; Pristionchus.  
Eukaryota; Metazoa; Chromadorea; Diplogasterida;  
Pristionchus pacificus

**TITLE** An integrated physical and genetic map of the nematode *Pristionchus pacificus*

**JOURNAL** Mol. Genet. Genomics 269 (5), 715-722 (2003)

**PUBMED** 12884007

COMMENT

Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
Class: BAC ends.

```

FEATURES
source
1. .1210
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol type="genomic DNA"

```

```
name="PpA EcoRI BAC Library"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="PpA EcoRI BAC Library"
/note="The library was generated by a partial digest of
```

ORIGIN

## ORIGIN

```
Query Match      9.6%; Score 83.4; DB 12; Length 1210;
Best Local Similarity 40.6%; Pred. No. 0.0003;
Matches 354; Conservative 0; Mismatches 516; Indels 2; Gaps 1;

QY 4 AAAAAAGTTATATTGCTGGAAATGGACCAAGTTTAAAGAAATGATATTCAAGGCTA 63
   |||||
Db 995 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 936
   |||||

QY 64 CCAATGATTTGATGTTAGATGTAATCAATTTTATTTTGAAGATAAATACTATCTT 123
   |||||
Db 935 TTAATTTATTTTAAATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 876
   |||||

QY 124 GGTAAAAATCAAGCAGTGTTTTACACCCCTGGTTCCTTTTGAGCAATACTACACT 183
   |||||
Db 875 ATTATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 816
   |||||

QY 184 TTAACAATTTAAATCCAAATCAAGATATGAGACCGAATCAATTTATGTTCTAATTAC 243
   |||||
Db 815 AAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 756
   |||||

QY 244 AACCAAGCTCATCTAGAAATGAAATTTTGTAAATCTTTTACGATTTTTCCTGAT 303
   |||||
Db 755 AATTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 696
   |||||

QY 304 GCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAATTT 363
   |||||
Db 695 AATTATTTAAATTTTAAATTTTATATATAATTAATAATANAATAAATAATTTATTT 636
   |||||

QY 364 CACGAATTTATTTCAA--TCAAAGAAATCTCTCAGGGGTCTATATGTCGAGTAGCCA 421
   |||||
Db 635 TTTNAAAAATTTATATATATTTTATTAANAATTTAAATTTTAAATTTTAAATTTAA 576
   |||||

QY 422 TAGCCCTAGGATACAAAGAAATTTATCTTCCTCGGAATTTGATTTTATCAATCGGTCTAT 481
   |||||
Db 575 ANTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 516
   |||||

QY 482 CTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATG 541
   |||||
Db 515 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 456
   |||||

QY 542 ATCGCTCACATATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTGAATTTCTAG 601
   |||||
Db 455 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 396
   |||||

QY 602 AAAAAACTTACAAATATAAATCTATATGCTTATGTCCTTAACAGTCTTTTACGAAATTTTA 661
   |||||
Db 395 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 336
   |||||

QY 662 TAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAATAAATACTACACTA 721
   |||||
Db 335 NAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 276
   |||||

QY 722 AAGATATATCTCATACCTCTTAGTGAGGCTTATGAAATTTTCAAAAAATATTTAATTTTA 781
   |||||
Db 275 NNNAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 216
   |||||

QY 782 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 841
   |||||
Db 215 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 156
   |||||

QY 842 CTAGTGATATAAGCATTTATTTTCAAGGAAAA 873
   |||||
Db 155 NNANNAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 124
   |||||
```

```
RESULT 13
CL508886
LOCUS CL508886 1258 bp DNA linear GSS 01-APR-2004
DEFINITION SAIL_804_E01.v1 SAIL Collection Arabidopsis thaliana genomic clone
ACCESSION SAIL_804_E01.v1, Genomic survey sequence.
VERSION CL508886
CL508886.1 GI:46006206
```

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## PUBMED

## COMMENT

## GSS.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicaceae; Arabidopsids.

1 (bases 1 to 1258)

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S. A.

A high-throughput Arabidopsis reverse genetics system

Plant Cell 14 (12), 2985-2994 (2002)

12468722

Contact: Sessions A

Applied Trait Genetics

Syngenta Biotechnology Inc.

3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA

Email: allen.sessions@syngenta.com

ABRC Stock Number C8935931; T-DNA left border flanking sequences of

Syngenta Arabidopsis Insertion Library (SAIL) lines are available

through the Arabidopsis Biological Resource Center (ABRC).

Sequences represent a pool of amplified genomic regions and not

single contiguous sequences.

Class: TDNA tagged.

Location/Qualifiers

## source

1. 1258

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/ecotype="Columbia"

/db\_xref="taxon:3702"

/clone="SAIL 804 E01.v1"

/clone\_lib="SAIL\_Collection"

/note="T-DNA left border sequences were isolated using a

modified TAIL-PCR strategy"

## ORIGIN

## Query Match

Best Local Similarity 9.5%; Score 83; DB 13; Length 1258;

Matches 348; Conservative 0; Mismatches 504; Indels 2; Gaps 2;

QY 1 ATGAAAAAGTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGG 60

Db 397 ATNNNNNNNNNNATTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 456

QY 61 CTACCAATGATTTTGATGCTATTAGATGCTATCAATTTTATTTTGAAGATAAATACTAT 120

Db 457 TTTTATTTATATANTATTTTANTATTTATATATATTTTATTTTAAATTTTATTTTAT 516

QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCCTTTGAGCAATACTAC 180

Db 517 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 576

QY 181 ACTTTAAACAATTTAATCCAAATCAAGAAATGAGACCGAACTAATATGTTCTTAAT 240

Db 577 AAATTTATANTTTTNTTAAATTTATTTTATATAATTTTATATAATTTTAAATAATATA 636

QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTTTCCT 300

Db 637 TAATAATATATTTATTTTATATAAATTTTNTAATTTTNTAATTTAATNTAATAAATTTTA 696

QY 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTTAA 360

Db 697 TATANATATATTTTATATAATATAATTTTAAATAAATAAATAAATAAATAAATAA 756

QY 361 TTTCAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Db 757 TATATATTTTAAATTTTATTTTATTTTAAATAAATAAATAAATAAATAAATAAATA 816

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480

Db 817 NNAANNATTTATTAATAAATAAATTTAATANTNTAATNTAATAAATAAATAAATAAATA 876

QY 481 TCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAAACTAGCCCTCGATTTTAAAT 540  
Db 877 TTAATTAANTTAATTAATATAA-TTTAAATTAATAAATAAATAAATAAATAAATAAT 935  
QY 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
Db 936 AAAATATAATATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 995  
QY 601 GAAATAACTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 660  
Db 996 ATANATANATAATATTNATATATATTNATATATTNATATATTNATATATAAATAAATAAATAAAT 1055  
QY 661 ATAGAACTAGCCGCAAAATTTAAATTTTAAATTTTATCATACAGAAAAATAAATAAATAAAT 720  
Db 1056 ATATATATATNAAATAAANNNTTANAANTTANAATAAATAAATAAATAAATAAATAAAT 1115  
QY 721 AAAGATATACTCATACCTTCTAGTGGCTTATGGAATTTTCAAAAAATAAATAAATAAAT 780  
Db 1116 NTATTAATTNATAATTNTTAATA-NNAATAAATAAATAAATAAATAAATAAATAAATAAAT 1174  
QY 781 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 840  
Db 1175 ATNANATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1234  
QY 841 CCTAGTGATATAA 854  
Db 1235 AATANNNTANTANA 1248

RESULT 14  
AG347098/c  
LOCUS AG347098 1268 bp DNA linear GSS 18-DEC-2004  
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-142102.T7, genomic survey sequence.  
ACCESSION AG347098  
VERSION AG347098.1 GI:47920408  
KEYWORDS GSS.  
SOURCE Mus musculus molossinus (Japanese wild mouse)  
ORGANISM Mus musculus molossinus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriaki, K. and Shiroishi, T.  
Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis  
Genome Res. 14 (12), 2439-2447 (2004)  
15574823

JOURNAL PUBMED  
REFERENCE 2 (bases 1 to 1268)  
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.  
Direct Submission  
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:hattori@ipc.riken.jp, URL:http://hgpc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp).  
Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@rtc.riken.jp

PRIMERS  
LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EcORI  
R.Site 2 : EcoRI.  
Location/Qualifiers  
COMMENT

source 1. 1268  
/organism="Mus musculus molossinus"  
/mol\_type="genomic DNA"  
/sub\_species="molossinus"  
/db\_xref="taxon:57486"  
/clone="MSMg01-142102.T7"  
/sex="male"  
/tissue\_types="mixture of kidney and spleen"  
/clone\_lib="MSMg01 Mouse Male BAC Library"  
ORIGIN  
Query Match 9.5%; Score 82.8; DB 14; Length 1268;  
Best Local Similarity 47.1%; Pred. No. 0.00038;  
Matches 321; Conservative 0; Mismatches 354; Indels 7; Gaps 2;  
QY 184 TTAATAACATTTAATCCAAATCAAGATATGAGACCACTAAATATGTGTCTTAATAC 243  
Db 1099 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1040  
QY 244 AACCAAGCTCATCTAGAGAAATGAAATTTTGTAAAACTTTTTCAGATTTTTCCTGAT 303  
Db 1039 TTATAATAAAATATAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAAT 980  
QY 304 GCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAAGTCTTATTTTAAATTT 363  
Db 979 AATTATTTTAAAAATAATATTTTATAAAAAATAATTTTAAATATAAAAAATAAATAA 920  
QY 364 CACGAAATTTTTCATCAAGAAATTTACCTCAGGGCTCTATATGTGCGAGTAGCCATA 423  
Db 919 TATAATAATTAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 860  
QY 424 GCCTAGGATACAAAGAAATTTATCTTTCGGGATTTGATTTTATCAAAATGGTCACT 483  
Db 859 AAAATTAATAATAAATAAATTTTATTTTATATAAATTAATAAATAAATAAATAAATAA 800  
QY 484 TATGCTTTTGATACCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAATAAGAT 543  
Db 799 AATAATAAAAAATATTAATATAATAATAAATAAATAAATAAATAAATAAATAAATAA 740  
QY 544 CGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTAG 603  
Db 739 AATAATAAATAAATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 682  
QY 604 AAAACTTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 663  
Db 681 AATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622  
QY 664 GAACCTAGCCCAAAATTTAAATTTTCAAAATTTTATCATACAGAAAAATAAATAAATAA 723  
Db 621 AAAATAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 562  
QY 724 GATATACTCATACCTTCTAGTGGCTTAT- - - - -GGAAAAATTTTCAAAAAATAAAT 778  
Db 561 AAAAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 502  
QY 779 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 838  
Db 501 TAAATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 442  
QY 839 TACCTAGTGATATAAAGCATTA 860  
Db 441 AATAAGAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA

RESULT 15  
AG376784/c  
LOCUS AG376784 1324 bp DNA linear GSS 21-DEC-2004  
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-184109.T7, genomic survey sequence.  
ACCESSION AG376784  
VERSION AG376784.1 GI:47987989  
KEYWORDS GSS.  
SOURCE Mus musculus molossinus (Japanese wild mouse)







```
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAATATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAAGCTAATATATGTGTTCTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATATTTTCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360
QY 361 TTTTCAAGAAATTTTAAATCAAGAAATTTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTAAATCAAGAAATTTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATGCTCAGCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAATTTCTA 600
Db 541 GATGCTCAGCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAATTTCTA 600
QY 601 GAAAAAATTTTACAAAATAAATACTATATGCTTATGCTTACAGCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTTACAAAATAAATACTATATGCTTATGCTTACAGCTTTTACGAAATTTT 660
QY 661 ATAGAATCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAATAAATCTACACT 720
Db 661 ATAGAATCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAATAAATCTACACT 720
QY 721 AAAGATATCTATACCTTTAGTGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATACCTTTAGTGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
QY 781 AAAAAATAAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAAGCATTTTCAAGGAAAA 873
```

## RESULT 2

```
US-10-303-162-11
; Sequence 11, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
```

```
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-162-11
```

```
Query Match 100.0%; Score 873; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 9e-162;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAGTTTATTTATGCTCGAAATGAGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTTATTTATGCTCGAAATGAGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
QY 61 CTACCAAAATGATTTTGTATTTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGTATTTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120
QY 121 CTTTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCCTTTTGAGCAATACTAC 180
Db 121 CTTTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCCTTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAAGCTAATATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAAGCTAATATATGTGTTCTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTTTCCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTTTCCT 300
QY 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTTATTTAAA 360
QY 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGTCTATATGTGTCAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATGCTCAGCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
Db 541 GATGCTCAGCTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
QY 601 GAAAAAATTTTACAAAATAAATACTATATGCTTATGCTTACAGCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTTACAAAATAAATACTATATGCTTATGCTTACAGCTTTTACGAAATTTT 660
QY 661 ATAGAATCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAATAAATCTACACT 720
Db 661 ATAGAATCTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAATAAATCTACACT 720
QY 721 AAAGATATCTATACCTTTAGTGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
Db 721 AAAGATATCTATACCTTTAGTGGCTTTATGGAATTTTCAAAATTTTAAATTTT 780
QY 781 AAAAAATAAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAAGCATTTTCAAGGAAAA 873
```

```
RESULT 3
US-10-303-134-11
; Sequence 11, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-134-11

Query Match 100.0%; Score 873; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 9e-162;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTAATCAAGG 60
DB 1 ATGAAAAAGTTATTATTCTGGAATGACCAAGTTTAAAGAAATGATTAATCAAGG 60
QY 61 CTACCAAAATGATTTGATGTTATGATGTAATCAATTTTATTTGCAAGTAATACTAT 120
DB 61 CTACCAAAATGATTTGATGTTATGATGTAATCAATTTTATTTGCAAGTAATACTAT 120
QY 121 CTGTGTAATAATGCAAGCAAGTTTACACCCCTGGTTCTTCTTGAGCAATACTAC 180
DB 121 CTGTGTAATAATGCAAGCAAGTTTACACCCCTGGTTCTTCTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCAAAATCAAGAAATGACCGAACTAATTAATGTTCTAAT 240
DB 181 ACTTTAAACATTTAATCAAAATCAAGAAATGACCGAACTAATTAATGTTCTAAT 240
QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTCT 300
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTAATTTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAA 360
DB 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAA 360
QY 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTACC 420
DB 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTACC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
QY 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600
```

```
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 600
QY 601 GAAAAAACTTTACAAAAATAAACTATATTGCTTTATGCTCTTAAACAGTCTTTTACGAAATTTT 660
DB 601 GAAAAAACTTTACAAAAATAAACTATATTGCTTTATGCTCTTAAACAGTCTTTTACGAAATTTT 660
QY 661 ATGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAAATCACTACT 720
DB 661 ATGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAAATCACTACT 720
QY 721 AAGATATATCTCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTAATTTT 780
DB 721 AAGATATATCTCATACCTTCTAGTGAGGCTTTATGGAATTTTCAAAAAATTAATTTT 780
QY 781 AAAAAATAAAAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
DB 781 AAAAAATAAAAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAAGCAATTTTCAAAAGGAAAA 873
DB 841 CCTAGTGATATAAAGCAATTTTCAAAAGGAAAA 873
```

## RESULT 4

```
US-10-303-118-11
; Sequence 11, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-118-11
```

```
Query Match 100.0%; Score 873; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 9e-162;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTAATCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTAATCAAGG 60
QY 61 CTACCAAAATGATTTGATGTTATGATGTAATCAATTTTATTTGAGATAAATACTAT 120
DB 61 CTACCAAAATGATTTGATGTTATGATGTAATCAATTTTATTTGAGATAAATACTAT 120
QY 121 CTGTGTAATAATGCAAGCAAGTTTACACCCCTGGTTCTTCTTGAGCAATACTAC 180
DB 121 CTGTGTAATAATGCAAGCAAGTTTACACCCCTGGTTCTTCTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCAAAATCAAGAAATGACCGAACTAATTAATGTTCTAAT 240
```

Db	181	ACTTTAAACATTTAATCCAAAATCAGAAATAGAGCCGAACTAAATTATGTGTTCTAAT	240
Qy	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTFAAAAACTTTTACGATTAATTTTCCT	300
Db	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTGTFAAAAACTTTTACGATTAATTTTCCT	300
Qy	301	GATGCTCATTTGGGATATGATTTTTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA	360
Db	301	GATGCTCATTTGGGATATGATTTTTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTAAA	360
Qy	361	TTTTCAACGAAATTTATTTCAATCAAAAGAAATTAACCTCAGGGTCTATATGTGTGCGATGACC	420
Db	361	TTTTCAACGAAATTTATTTCAATCAAAAGAAATTAACCTCAGGGTCTATATGTGTGCGATGACC	420
Qy	421	ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAAATGATTTTTTATCAAAAATGGGTCA	480
Db	421	ATAGCCCTTAGGATACAAAGAAATTTATCTTTTCGGGAAATGATTTTTTATCAAAAATGGGTCA	480
Qy	481	TCATTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTTGATTTTAAAAAAT	540
Db	481	TCATTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTTGATTTTAAAAAAT	540
Qy	541	GATCGCTCACACTATATCGGACATAGTAAAAAATACAGATATAAAGCTTTAGAAATTTCTA	600
Db	541	GATCGCTCACACTATATCGGACATAGTAAAAAATACAGATATAAAGCTTTAGAAATTTCTA	600
Qy	601	GAIAAAACTTTACAAAATAAAACTATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTTT	660
Db	601	GAIAAAACTTTACAAAATAAAACTATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTTT	660
Qy	661	ATAGAACTAGCGCCAAATTTTAAATTCAAATTTTATCATACAAGAAAAAAATAAATTAACACT	720
Db	661	ATAGAACTAGCGCCAAATTTTAAATTCAAATTTTATCATACAAGAAAAAAATAAATTAACACT	720
Qy	721	AAAGATATACCTCATACCTCTCTAGTGAGCTTATCGGAAATTTTCAAAAAATTAATTAATTT	780
Db	721	AAAGATATACCTCATACCTCTCTAGTGAGCTTATCGGAAATTTTCAAAAAATTAATTAATTT	780
Qy	781	AAAAAAATAAAAAATTTAAAGAAAAATGTTTTATTACAAGTTTGAATAAAGATCTATTAAAGATTA	840
Db	781	AAAAAAATAAAAAATTTAAAGAAAAATGTTTTATTACAAGTTTGAATAAAGATCTATTAAAGATTA	840
Qy	841	CCTAGTGATATTAAGCATTAATTTCAAAGGAAAA	873
Db	841	CCTAGTGATATTAAGCATTAATTTCAAAGGAAAA	873

RESULT 5  
US-10-303-128-11  
; Sequence 11, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-128-11

Query Match 100.0%; Score 873; DB 3; Length 873;
Best Local Similarity 100.0%; Pred. No. 9e-162;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTCTCGAAATGACCAAGTTTAAAGAAATTCATTATTCACAGG 60
   |||||
Db 1 ATGAAAAAGTTATTCTCGAAATGACCAAGTTTAAAGAAATTCATTATTCACAGG 60

QY 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120
   |||||
Db 61 CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGAAAGATAAATACTAT 120

QY 121 CTTGGTAAATAATGCAAAAGCAGTGTTTTACCCCTGGTTCCTTTTGAGCAATACTAC 180
   |||||
Db 121 CTTGGTAAATAATGCAAAAGCAGTGTTTTACCCCTGGTTCCTTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACTAATATGTGTTCTAAT 240
   |||||
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACTAATATGTGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCC 300
   |||||
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCC 300

QY 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTATTTTAAA 360
   |||||
Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTATTTTAAA 360

QY 361 TTTTCAGCAATTTATTTTCAATCAAGAAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420
   |||||
Db 361 TTTTCAGCAATTTATTTTCAATCAAGAAATTTACTCAGGGGTCTATATGTGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480
   |||||
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
   |||||
Db 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600
   |||||
Db 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTA 600

QY 601 GAAAAAACCTTACAAAATAAAAACTATATGCTTTATGCTCTTAAACAGATCTTTTAGCAATTTT 660
   |||||
Db 601 GAAAAAACCTTACAAAATAAAAACTATATGCTTTATGCTCTTAAACAGATCTTTTAGCAATTTT 660

QY 661 ATAGAACTAGCGGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATACTACACT 720
   |||||
Db 661 ATAGAACTAGCGGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATACTACACT 720

QY 721 AAGATATATCTCATCTTCTAGTCAGGCTTATGGAATTTTCAAAAATAATTAATTTT 780
   |||||
Db 721 AAGATATATCTCATCTTCTAGTCAGGCTTATGGAATTTTCAAAAATAATTAATTTT 780

QY 781 AAAAAAATAAAAAATTAAGAAATTTGTTTATTAACAAGTTGATAAAAAGATCTATAAGATTA 840
   |||||
Db 781 AAAAAAATAAAAAATTAAGAAATTTGTTTATTAACAAGTTGATAAAAAGATCTATAAGATTA 840

QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAA 873
   |||||
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAA 873

```

## RESULT 6

US-10-735-419-11  
; Sequence 11, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-735-419-11

Query Match 100.0%; Score 873; DB 5; Length 873;  
Best Local Similarity 100.0%; Pred. No. 9e-162;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
DB 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAAAATGATTTTCATGATTTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120  
DB 61 CTACCAAAATGATTTTCATGATTTAGATGTAATCAATTTTATTTGAAGATAAATCTAT 120

QY 121 CTGTGTAATAATGCAAGCAGTGTGTTACACCCCTGGTTCCTCTTTCAGCAATACTAC 180  
DB 121 CTGTGTAATAATGCAAGCAGTGTGTTACACCCCTGGTTCCTCTTTCAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACAAATATATGTTCTAAT 240  
DB 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACAAATATATGTTCTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACGATTATTTCT 300  
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACACTTTTACGATTATTTCT 300

QY 301 GATGCTCATTTGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTTATTTAA 360  
DB 301 GATGCTCATTTGGATATGATTTTAAACAACTTAAAGAAATTTAAAGCTTTATTTAA 360

QY 361 TTTCAGCAATTTATTTCAATCAAGAATTAATCCTCAGGGGTCTATATGTTGTCAGTAC 420  
DB 361 TTTCAGCAATTTATTTCAATCAAGAATTAATCCTCAGGGGTCTATATGTTGTCAGTAC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTATACCAAAACAAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAT 540  
DB 481 TCTTATGCTTTTATACCAAAACAAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCACATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTTGAATTTCTA 600

```

Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGTAAGAACTTTTACGATATTTTCT 300
Qy 301 GATCCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAAGCTTATTTAA 360
Db 301 GATCCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAAGCTTATTTAA 360
Qy 361 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATATGTGCGAGTAGCC 420
Db 361 TTTCACGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATATGTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCACACTATATCGGCATAGTAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTCAACAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAACTATATGCTTATGCTCAACAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAAATTAACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAAATTAACACT 720
Qy 721 AAAGATATATCTATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780
Db 721 AAAGATATATCTATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780
Qy 781 AAAAAATAAAAAATTAAGAAATTTTATTTATCAAGTTGATATAAGATTTTAAAGATTA 840
Db 781 AAAAAATAAAAAATTAAGAAATTTTATTTATCAAGTTGATATAAGATTTTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

```

```

RESULT 8
US-10-303-162-13
; Sequence 13, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303.162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)

```

```

; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-162-13
Query Match 99.3%; Score 866.6; DB 3; Length 873;
Best Local Similarity 99.5%; Pred. No. 1.6e-160;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTATTTATTTGCTGGAATGAGCAAGCTTTAAAAAGAAATTCGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTTATTTGCTGGAATGAGCAAGCTTTAAAAAGAAATTCGATTTTCAAGG 60
Qy 61 CTACCAATGATTTTATGATGATTTATGATGATCAATTTTATTTGAAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTATGATGATTTATGATGATCAATTTTATTTTCAAGATAAATACTAT 120
Qy 121 CTGCTGTAATAATGCAAAAGCAGTGTTTTACACCCCTGTTTCTTTGAGCAATACTAC 180
Db 121 CTGCTGTAATAATGCAAAAGCAGTGTTTTACACCCCTGTTTCTTTGAGCAATACTAC 180
Qy 181 ACTTTAAAAACATTTTAAATCAAAATCAAGAAATGAGCCGAATCAATTTATGTTCTTAAT 240
Db 181 ACTTTAAAAACATTTTAAATCAAAATCAAGAAATGAGCCGAATCAATTTATGTTCTTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTTCCT 300
Qy 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAACTGTTATTTTAAA 360
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTTAACTGTTATTTTAAA 360
Qy 361 TTTTACGAAATTTTATTTCAATCAAAAGAAATTTTCTCAGGGCTCTATATGTTGCGAGTAGCC 420
Db 361 TTTTACGAAATTTTATTTCAATCAAAAGAAATTTTCTCAGGGCTCTATATGTTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCACACTATATCGGCATAGTAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAAATACAGATATAAAAGCTTTTGAATTTCTA 600
Qy 601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAAGCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAAGCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAAATTAACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAAATTAACACT 720
Qy 721 AAAGATATATCTATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780
Db 721 AAAGATATATCTATACCTTTCTAGTGAGGCTTTATGGAATTTTCAAAAATTTTAAATTT 780
Qy 781 AAAAAATAAAAAATTAAGAAATTTTATTTATCAAGTTGATATAAGATTTTAAAGATTA 840
Db 781 AAAAAATAAAAAATTAAGAAATTTTATTTATCAAGTTGATATAAGATTTTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAA 873

```

```

RESULT 9
US-10-303-134-13
; Sequence 13, Application US/10303134
; Patent No. 6825019

```

```

; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
; US-10-303-134-13

Query Match          99.3%; Score 866.6; DB 3; Length 873;
Best Local Similarity 99.5%; Pred. No. 1.6e-160;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGCTTTAAAGAAATTTGATTATTTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGCTTTAAAGAAATTTGATTATTTCAAGG 60

QY 61 CTACCAATGATTTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120
DB 61 CTACCAATGATTTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120

QY 121 CTTCGTAATAATGCAAGCAGTGTTCACACCCCTGGTTCCTTTTGGAGCAATCTAT 180
DB 121 CTTCGTAATAATGCAAGCAGTGTTCACACCCCTGGTTCCTTTTGGAGCAATCTAT 180

QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAACTAATTTATGTTCTTAAT 240
DB 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAACTAATTTATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

QY 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGNGCGAGTAGCC 420
DB 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGNGCGAGTAGCC 420

QY 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCATTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
DB 481 TCATTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCACACTATATCGGACATAGTAAATACAGATATAGAAAGCTTTTGAATTTCTA 600
DB 541 GATCGCTCACACTATATCGGACATAGTAAATACAGATATAGAAAGCTTTTGAATTTCTA 600

QY 601 GAAAAAACTTACAAAATAAAACTATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTT 660

```

```

DB 601 GAAAAAACTTACAAAATAAAACTATATGCTTATGTCCTTAATAGTCTTTTAGCAAAATTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720
QY 721 AAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 780
DB 721 AAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 780
QY 781 AAAAAAATAAATAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTA 840
DB 781 AAAAAAATAAATAATTAAGAAATGTTTATTAACAAGTTGATAAAGATCTTATTAAGATTA 840
QY 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
DB 841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873

RESULT 10
US-10-303-118-13
; Sequence 13, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
; US-10-303-118-13

Query Match          99.3%; Score 866.6; DB 3; Length 873;
Best Local Similarity 99.5%; Pred. No. 1.6e-160;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTCTGCGAAATGGACCAAGCTTTAAAGAAATTTGATTATTTCAAGG 60
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGCTTTAAAGAAATTTGATTATTTCAAGG 60

QY 61 CTACCAATGATTTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120
DB 61 CTACCAATGATTTTGATGTTATAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120

QY 121 CTTCGTAATAATGCAAGCAGTGTTCACACCCCTGGTTCCTTTTGGAGCAATCTAT 180
DB 121 CTTCGTAATAATGCAAGCAGTGTTCACACCCCTGGTTCCTTTTGGAGCAATCTAT 180

QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAACTAATTTATGTTCTTAAT 240
DB 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGACCGAACTAATTTATGTTCTTAAT 240

QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

QY 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGNGCGAGTAGCC 420
DB 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGNGCGAGTAGCC 420

QY 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
DB 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480

QY 481 TCATTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540
DB 481 TCATTATGCTTTTGATACAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCACACTATATCGGACATAGTAAATACAGATATAGAAAGCTTTTGAATTTCTA 600
DB 541 GATCGCTCACACTATATCGGACATAGTAAATACAGATATAGAAAGCTTTTGAATTTCTA 600

QY 601 GAAAAAACTTACAAAATAAAACTATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTT 660

```



```

Db 241 TACACCAAGCTCATCTAGAAAATGAAATTTTGTAAACCTTTTACGATTATTTTCCCT 300
Qy 301 GATGCTCATTTGGGATGATGTTTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATGATGTTTTTAAACCAACTTAAAGAAATTAATGCTTTATTTAAA 360
Qy 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATATGTCAGTAGCC 420
Db 361 TTTTACGAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATATGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA 600
Qy 601 GAAAAAATTTACAAATAAATACTATATGCTTATGCTTACAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACAAATAAATACTATATGCTTATGCTTACAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCCCAAAATTTAAATTTTAAATTTTATCATACAAAGAAATAAATCTACT 720
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTTAAATTTTATCATACAAAGAAATAAATCTACT 720
Qy 721 AAAGATATCTACTACTCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAATAATTTT 780
Db 721 AAAGATATCTACTACTCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAATAATTTT 780
Qy 781 AAAAAATAAATAAATAAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAATAAATAAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873

```

```

RESULT 11
US-10-303-128-13
; Sequence 13, Application US/10303128
; Patent No. 691337
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II

```

```

; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-128-13
Query Match 99.3%; Score 866.6; DB 3; Length 873;
Best Local Similarity 99.5%; Pred. No. 1.6e-160;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAAAAGTTTATTTATGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTATTTCAAGG 60
Db 1 ATGAAAAAGTTTATTTATGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTATTTCAAGG 60
Qy 61 CTACCAATGATTTTATGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Db 61 CTACCAATGATTTTATGATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120
Qy 121 CTGTGTAATAAATGCAAAAGCAGTGTTTACACCCCTGGTTTCTTCTTTCAGCAATACTAC 180
Db 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTAAATTTCTTCTTTCAGCAATACTAC 180
Qy 181 ACTTTAAACATTTTAAATCCAAATCAAGATATAGACCGAACTAAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATAGACCGAACTAAATATGTGTTCTAAT 240
Qy 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTGAAAAAATTTTACGATTATTTTCCCT 300
Db 241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTGAAAAAATTTTACGATTATTTTCCCT 300
Qy 301 GATGCTCATTTTGGGATGATGTTTTTAAACCAACTTAAAGAAATTTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTTGGGATGATGTTTTTAAACCAACTTAAAGAAATTTTAAATGCTTTATTTAAA 360
Qy 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGGCTATATATGTGTCAGTAGCC 420
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGGCTATATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Qy 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTTAGAATTTCTA 600
Qy 601 GAAAAAATTTACAAATAAATACTATATGCTTATGCTTAAACAGTCTTTTACGAAATTTT 660
Db 601 GAAAAAATTTACAAATAAATACTATATGCTTATGCTTAAACAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCCCAAAATTTAAATTTTAAATTTTATCATACAAAGAAATAAATCTACT 720
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTTAAATTTTATCATACAAAGAAATAAATCTACT 720
Qy 721 AAAGATATCTACTACTCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAATAATTTT 780
Db 721 AAAGATATCTACTACTCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAATAATTTT 780
Qy 781 AAAAAATAAATAAATAAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAATAAATAAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873

```

```

RESULT 12
US-10-735-419-13
; Sequence 13, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:

```



APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-00011105  
CURRENT APPLICATION NUMBER: US/10/735,419  
CURRENT FILING DATE: 2003-12-11  
PRIOR APPLICATION NUMBER: US/09/816,028A  
PRIOR FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
PRIOR APPLICATION NUMBER: US 09/495,406  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 13  
LENGTH: 873  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(873)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-735-419-13

Query Match 99.3%; Score 866.6; DB 5; Length 873;  
Best Local Similarity 99.5%; Pred. No. 1.6e-160;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
QY 61 CTACCAATGATTTTATGATTTAGATGATCAATTTTATTTGAAGATAAATCTAT 120  
Db 61 CTACCAATGATTTTATGATTTAGATGATCAATTTTATTTGAAGATAAATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTGTTCTCTTTGAGCAATCTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTGTTCTCTTTGAGCAATCTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 660

Db 601 GAAAAAAGTTACAAAAATAAACTATATGCTTATGCTCTATAGTCTTTTAGCAAAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTTAAATTTCAATTTTATCATACAAGAAAAATAACTACTACT 720  
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTCAATTTTATCATACAAGAAAAATAACTACTACT 720  
QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 781 AAAAAAATAAATAATTAAGAAAAATGTTTATTTACAAGTTGATAAAGATCTATTAGATTA 840  
Db 781 AAAAAAATAAATAATTAAGAAAAATGTTTATTTACAAGTTGATAAAGATCTATTAGATTA 840  
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 13  
US-09-495-406-8  
Sequence 8, Application US/09495406  
Patent No. 6503744  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
FILE REFERENCE: 019633-00011105  
CURRENT APPLICATION NUMBER: US/09/495,406  
CURRENT FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: US 60/118,213  
PRIOR FILING DATE: 1999-02-01  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 8  
LENGTH: 876  
TYPE: DNA  
ORGANISM: Campylobacter jejuni  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(876)  
OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
OTHER INFORMATION: (CstII) from C. jejuni O:19  
US-09-495-406-8

Query Match 99.3%; Score 866.6; DB 3; Length 876;  
Best Local Similarity 99.5%; Pred. No. 1.6e-160;  
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
QY 61 CTACCAATGATTTTATGATTTATGATGATTAATCAATTTTATTTGAAGATAAATCTAT 120  
Db 61 CTACCAATGATTTTATGATTTATGATGATTAATCAATTTTATTTGAAGATAAATCTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTGTTCTCTTTGAGCAATCTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTGTTCTCTTTGAGCAATCTAC 180  
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGACCGAATTAATATGTTCTTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360

Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
 Qy 361 TTTTACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGC 420  
 Db 361 TTTTACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGC 420  
 Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAAATGGGTCA 480  
 Qy 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAGAAAT 540  
 Db 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAGAAAT 540  
 Qy 541 GATCGCTCAGCTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTTAAAGAAATTTCTA 600  
 Db 541 GATCGCTCAGCTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTTAAAGAAATTTCTA 600  
 Qy 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660  
 Db 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660  
 Qy 661 ATAGAACTAGCGCCAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 720  
 Db 661 ATAGAACTAGCGCCAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 720  
 Qy 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTTATGGAATTTTAAAGAAATTTTAAAGAAATTTT 780  
 Db 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTTATGGAATTTTAAAGAAATTTTAAAGAAATTTT 780  
 Qy 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840  
 Db 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840  
 Qy 841 CCTAGTGATATAAGCAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 873  
 Db 841 CCTAGTGATATAAGCAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 873

RESULT 14

US-09-816-028A-8  
 ; Sequence 8, Application US/09816028A  
 ; Patent No. 6699705  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/09/816,028A  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 ; US-09-816-028A-8

Query Match 99.3%; Score 866.6; DB 3; Length 876;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-160;  
 Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTTATGCTGGAATGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60  
 Db 1 ATGAAAAAGTTATTTATGCTGGAATGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60  
 Qy 61 CTACCAATGATTTTATGATGATTTATGATGATTTATGATGATTTATTTTGAAGATATAACTAT 120  
 Db 61 CTACCAATGATTTTATGATGATTTATGATGATTTATGATGATTTATTTTGAAGATATAACTAT 120  
 Qy 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTGGTTTCTTTCTTTGAGCAATACTAC 180  
 Db 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTGGTTTCTTTCTTTGAGCAATACTAC 180  
 Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAAATATGTTCTTAAT 240  
 Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATGAGACCGAACTAAATATGTTCTTAAT 240  
 Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTTCCCT 300  
 Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTTATTTTCCCT 300  
 Qy 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
 Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360  
 Qy 361 TTTTACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGC 420  
 Db 361 TTTTACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGGC 420  
 Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAAATGGGTCA 480  
 Qy 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAGAAAT 540  
 Db 481 TCTTATGCTTTTATACCAAAACAAAGAAATTTTAAACCTAGCCCTGATTTTAAAGAAAT 540  
 Qy 541 GATCGCTCAGCTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTTAAAGAAATTTCTA 600  
 Db 541 GATCGCTCAGCTATATCGGACATAGTAAAGAAATACAGATATAAAGCTTTTAAAGAAATTTCTA 600  
 Qy 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660  
 Db 601 GAAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 660  
 Qy 661 ATAGAACTAGCGCCAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 720  
 Db 661 ATAGAACTAGCGCCAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 720  
 Qy 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTTATGGAATTTTAAAGAAATTTTAAAGAAATTTT 780  
 Db 721 AAAGATATACCTATACCTTTCTAGTGAGGCTTTATGGAATTTTAAAGAAATTTTAAAGAAATTTT 780  
 Qy 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840  
 Db 781 AAAAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 840  
 Qy 841 CCTAGTGATATAAGCAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 873  
 Db 841 CCTAGTGATATAAGCAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTTAAAGAAATTTT 873

RESULT 15

US-10-303-162-8  
 ; Sequence 8, Application US/10303162  
 ; Patent No. 6723545  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/303,162

```

; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-162-8

Query Match          99.3%; Score 866.6; DB 3; Length 876;
Best Local Similarity 99.5%; Pred. No. 1.6e-160;
Matches 869; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 ATGAAAAAGATTATTATTGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTCACAGG 60
DB      1 ATGAAAAAGATTATTATTGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTCACAGG 60

QY      61 CTACCAAAATGATTTTGATGATATTAGATGTAATCAATTTTATTTGAGATAAATACTAT 120
DB      61 CTACCAAAATGATTTTGATGATATTAGATGTAATCAATTTTATTTGAGATAAATACTAT 120

QY      121 CTTGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTCTTGAGCAATACTAC 180
DB      121 CTTGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180

QY      181 ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAACTAATATATGTGTTCTAAT 240
DB      181 ACTTTAAACATTTAATCAAAATCAAGAAATGAGACCGAACTAATATATGTGTTCTAAT 240

QY      241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300
DB      241 TACAAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTAATTTTCT 300

QY      301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTTAA 360
DB      301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAAATGCTTTATTTTAA 360

QY      361 TTTCACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTACCC 420
DB      361 TTTCACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTACCC 420

QY      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
DB      421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

QY      481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
DB      481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY      541 GATCGCTCACAATATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600
DB      541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTTAGAATTTCTA 600

QY      601 GAAAAAATCTTACAAATAAAACTATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTT 660
DB      601 GAAAAAATCTTACAAATAAAACTATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTT 660

QY      661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAAATAACTACACT 720
DB      661 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAAATAACTACACT 720

QY      721 AAAGATATACTACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
```

```

|||||
Db      721 AAAGATATACTACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAATTTT 780
QY      781 AAAAAAATAAAAATTTAAAGAAATGTTTATTACAAAGTTGATATAAGATCTTATTAAAGATTA 840
|||||
Db      781 AAAAAAATAAAAATTTAAAGAAATGTTTATTACAAAGTTGATATAAGATCTTATTAAAGATTA 840
|||||
QY      841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
|||||
Db      841 CCTAGTGATATAAAGCATTATTTCAAAGGAAAA 873
|||||
```

Search completed: July 19, 2006, 15:09:24  
Job time : 137.177 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:09:43 ; Search time 996.192 Seconds  
(without alignments)  
10768.108 Million cell updates/sec

Title: US-10-734-719-11

Perfect score: 873

Sequence: 1 atgaaaaagttatttgc.....agcattatttcaagaaaaa 873

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 18992170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 10: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 11: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 12: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 13: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 14: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 15: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10J\_PUBCOMB.seq.\*
- 16: /EMC\_Celerra\_SIDS3/ptodata/2/pubpna/US10K\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	873	100.0	873	3	US-09-816-028A-11
2	873	100.0	873	7	US-10-303-161-11
3	873	100.0	873	7	US-10-303-118-11
4	873	100.0	873	7	US-10-303-128-11
5	873	100.0	873	7	US-10-303-134-11
6	873	100.0	873	7	US-10-303-162-11
7	873	100.0	873	7	US-10-735-419-11
8	873	100.0	873	9	US-10-820-536-11
9	873	100.0	873	9	US-10-845-408-11
10	873	100.0	873	9	US-10-845-412-11
11	873	100.0	873	9	US-10-846-219-11
12	873	100.0	873	9	US-10-821-604-11
13	873	100.0	873	9	US-10-847-983-11
14	873	100.0	873	9	US-10-821-573-11
15	873	100.0	873	9	US-10-850-807-11
16	873	100.0	873	9	US-10-850-125-11
17	873	100.0	873	9	US-10-830-825-11

18	873	100.0	873	10	US-10-962-334-11	Sequence 11, Appl
19	873	100.0	873	10	US-10-830-997-11	Sequence 11, Appl
20	873	100.0	873	10	US-10-962-235-11	Sequence 11, Appl
21	873	100.0	873	10	US-10-961-882-11	Sequence 11, Appl
22	866.6	99.3	873	3	US-09-816-028A-13	Sequence 13, Appl
23	866.6	99.3	873	7	US-10-303-161-13	Sequence 13, Appl
24	866.6	99.3	873	7	US-10-303-118-13	Sequence 13, Appl
25	866.6	99.3	873	7	US-10-303-128-13	Sequence 13, Appl
26	866.6	99.3	873	7	US-10-303-134-13	Sequence 13, Appl
27	866.6	99.3	873	7	US-10-303-162-13	Sequence 13, Appl
28	866.6	99.3	873	9	US-10-735-419-13	Sequence 13, Appl
29	866.6	99.3	873	9	US-10-820-536-13	Sequence 13, Appl
30	866.6	99.3	873	9	US-10-845-408-13	Sequence 13, Appl
31	866.6	99.3	873	9	US-10-845-412-13	Sequence 13, Appl
32	866.6	99.3	873	9	US-10-846-219-13	Sequence 13, Appl
33	866.6	99.3	873	9	US-10-821-604-13	Sequence 13, Appl
34	866.6	99.3	873	9	US-10-847-983-13	Sequence 13, Appl
35	866.6	99.3	873	9	US-10-821-573-13	Sequence 13, Appl
36	866.6	99.3	873	9	US-10-850-807-13	Sequence 13, Appl
37	866.6	99.3	873	9	US-10-850-125-13	Sequence 13, Appl
38	866.6	99.3	873	9	US-10-830-825-13	Sequence 13, Appl
39	866.6	99.3	873	10	US-10-962-334-13	Sequence 13, Appl
40	866.6	99.3	873	10	US-10-830-997-13	Sequence 13, Appl
41	866.6	99.3	873	10	US-10-962-235-13	Sequence 13, Appl
42	866.6	99.3	873	10	US-10-961-882-13	Sequence 8, Appl
43	866.6	99.3	876	3	US-09-816-028A-8	Sequence 8, Appl
44	866.6	99.3	876	7	US-10-303-161-8	Sequence 8, Appl
45	866.6	99.3	876	7	US-10-303-118-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1

US-09-816-028A-11  
; Sequence 11, Application US/09816028A  
; Patent No. US2002042369A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816, 028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-09-816-028A-11

Query Match 100.0%; Score 873; DB 3; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTCTGGAATGGACCAAGTTAAAGAAATGATTATTCAAGG 60  
QY 61 CTACCAATGATTTTGTATGATGATTAATCAATTTTATTGAGATAAATCTAT 120

```
Db 61 CTACCAATGATTTTGATGTATTTAGATGTATCAATTTTATTTTGAAGATAAATACAT 120
Qy 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTGGTCTTCTTTGAGCAATAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTGGTCTTCTTTGAGCAATAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACTAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACTAATATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCC 300
Qy 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTTACAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTTACAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTTGAATTTCTA 600
Qy 601 GAAAACTTACAAAAATTAATGCTTATGCTTCAAGTCTTACAGTCTTTTACGAAATTT 660
Db 601 GAAAACTTACAAAAATTAATGCTTATGCTTCAAGTCTTACAGTCTTTTACGAAATTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATGCTTCAAGTCTTACAGTCTTTTAC 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATGCTTCAAGTCTTACAGTCTTTTAC 720
Qy 721 AAAGATATCTATACCTTTAGTGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780
Db 721 AAAGATATCTATACCTTTAGTGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
```

## RESULT 2

```
US-10-303-161-11
; Sequence 11, Application US/10303161
; Publication No. US20030148459A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
```

```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.11
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-161-11
```

```
Query Match 100.0%; Score 873; DB 7; Length 873;
Best Local Similarity 100.0%; Pred. No. 2,4e-138;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAAAAGTTATTTATGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60
Db 1 ATGAAAAAGTTATTTATGCTCGAAATGGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60
Qy 61 CTACCAATGATTTTGTGATGATTTTATGATGATCAATTTTATTTTGAAGATAAATACAT 120
Db 61 CTACCAATGATTTTGTGATGATTTTATGATGATCAATTTTATTTTGAAGATAAATACAT 120
Qy 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTGGTCTTCTTTGAGCAATAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTGGTCTTCTTTGAGCAATAC 180
Qy 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACTAATATGTGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACTAATATGTGTTCTAAT 240
Qy 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCC 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTTATTTTCC 300
Qy 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Db 301 GATGCTCATTTGGGATGATTTTTTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 361 TTTTACAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Db 361 TTTTACAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGGGTCA 480
Qy 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 541 GATGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTTGAATTTTCTA 600
Db 541 GATGCTCACACTATATCGGCATAGTAAAAATACAGATATATAAGCTTTTGAATTTTCTA 600
Qy 601 GAAAACTTACAAAAATTAATGCTTATGCTTCAAGTCTTACAGTCTTTTACGAAATTTT 660
Db 601 GAAAACTTACAAAAATTAATGCTTATGCTTCAAGTCTTACAGTCTTTTACGAAATTTT 660
Qy 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATGCTTCAAGTCTTACAGTCTTTTAC 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTAACTATATGCTTCAAGTCTTACAGTCTTTTAC 720
Qy 721 AAAGATATCTATACCTTTAGTGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780
Db 721 AAAGATATCTATACCTTTAGTGGCTTATGGAATTTTCAAAAAATTTTCAAAAAATTTT 780
Qy 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
```

QY 841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873  
Db 841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873

RESULT 3  
US-10-303-118-11  
; Sequence 11, Application US/10303118  
; Publication No. US20030157655A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-118-11

Query Match 100.0%; Score 873; DB 7; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTTATTCTGGAATGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTTATTCTGGAATGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60

QY 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120

QY 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTGTTCTCTTTCGACCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTGTTCTCTTTCGACCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCAACTAAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCAACTAAATATGTTCTTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300

QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATCTTATTTTAA 360  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATCTTATTTTAA 360

QY 361 TTTCAGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTCAGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATATCAAAACAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATATCAAAACAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATATAAAGCTTTTGAATTTCTA 600  
Db 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATATAAAGCTTTTGAATTTCTA 600

QY 601 GAAAAAACTTTACAAAAATAAACTATATGCTTATGCTTCTAAACAGTCTTTTAGCAAAATTT 660  
Db 601 GAAAAAACTTTACAAAAATAAACTATATGCTTATGCTTCTAAACAGTCTTTTAGCAAAATTT 660

QY 661 ATGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATTAACACT 720  
Db 661 ATGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATTAACACT 720

QY 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780

QY 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAGATTA 840  
Db 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAGATTA 840

QY 841 CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA 873  
Db 841 CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA 873

RESULT 4  
US-10-303-128-11  
; Sequence 11, Application US/10303128  
; Publication No. US20030157656A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-303-128-11

Query Match 100.0%; Score 873; DB 7; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTTATTCTGGAATGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60  
Db 1 ATGAAAAAGTTTATTCTGGAATGACCAAGTTTAAAAAGAAATGATTATTTCAAGG 60

QY 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
Db 61 CTACCAATGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 120

QY 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTGTTCTCTTTCGACCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTGTTTACACCCCTGTTCTCTTTCGACCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCAACTAAATATGTTCTTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCAACTAAATATGTTCTTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300

QY 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATCTTATTTTAA 360  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATCTTATTTTAA 360

QY 361 TTTCAGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTCAGTAGCC 420  
Db 361 TTTCAGAAATTTATTTCAATCAAGAAATTTACTCAGGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 480

```
QY 121 CTTGGTAAATAATGCAAGCAGTGTATTACACCCCTGGTTCCTTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAATAATGCAAGCAGTGTATTACACCCCTGGTTCCTTCTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAATATGTTCTAAT 240
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAA 360
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTACC 420
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTACC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAATCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAATCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATGCTCACACTATATCGGACATAGTAAATACAGATATAGAACTTTTAAAGCTTTTCTA 600
Db 541 GATGCTCACACTATATCGGACATAGTAAATACAGATATAGAACTTTTAAAGCTTTTCTA 600
QY 601 GAAAAAATTTTACAAATTAATAATTTTATGCTTATGCTTAAACAGCTTTTACGAAATTT 660
Db 601 GAAAAAATTTTACAAATTAATAATTTTATGCTTATGCTTAAACAGCTTTTACGAAATTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATGCTTATGCTTAAACAGCTTTTAC 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATGCTTATGCTTAAACAGCTTTTAC 720
QY 721 AAAGATATCTATACCTTCTAGTGAGGCTTTATGCTTAAACAGCTTTTACGAAATTT 780
Db 721 AAAGATATCTATACCTTCTAGTGAGGCTTTATGCTTAAACAGCTTTTACGAAATTT 780
QY 781 AAAAAATTAATAATTAAGAAATTTTATTTATCAAGTTGATAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATTTTATTTATCAAGTTGATAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
```

## RESULT 5

```
US-10-303-134-11
; Sequence 11, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
```

```
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3(alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-134-11
```

```
Query Match 100.0%; Score 873; DB 7; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.4e-136;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTTTCAAGG 60
QY 61 CTACCAAAATGATTTTGTATGATATTTAGATCTAATCAATTTTATTTGCAAGATAAATACTAT 120
Db 61 CTACCAAAATGATTTTGTATGATATTTAGATCTAATCAATTTTATTTGCAAGATAAATACTAT 120
QY 121 CTTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCCTTTGAGCAATACTAC 180
Db 121 CTTGTAAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCCTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACTAATATGTTCTAAT 240
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACTAATATGTTCTAAT 240
QY 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300
Db 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAACCTTTTACGATTTTCTCT 300
QY 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAA 360
Db 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAA 360
QY 361 TTTACGAAATTTAATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTACC 420
Db 361 TTTACGAAATTTAATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTACC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATGCTTATGCTTAAACAGCTTTTACGAAATTT 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATGCTTATGCTTAAACAGCTTTTACGAAATTT 480
QY 481 TCTTATGCTTTTGATACCAATCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAATCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATGCTCACACTATATCGGACATAGTAAATACAGATATAGAACTTTTAAAGCTTTTCTA 600
Db 541 GATGCTCACACTATATCGGACATAGTAAATACAGATATAGAACTTTTAAAGCTTTTCTA 600
QY 601 GAAAAAATTTTACAAATTAATAATTTTATGCTTATGCTTAAACAGCTTTTACGAAATTT 660
Db 601 GAAAAAATTTTACAAATTAATAATTTTATGCTTATGCTTAAACAGCTTTTACGAAATTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATGCTTATGCTTAAACAGCTTTTAC 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATGCTTATGCTTAAACAGCTTTTAC 720
QY 721 AAAGATATCTATACCTTCTAGTGAGGCTTTATGCTTAAACAGCTTTTACGAAATTT 780
Db 721 AAAGATATCTATACCTTCTAGTGAGGCTTTATGCTTAAACAGCTTTTACGAAATTT 780
QY 781 AAAAAATTAATAATTAAGAAATTTTATTTATCAAGTTGATAAGATCTATTAAAGATTA 840
Db 781 AAAAAATTAATAATTAAGAAATTTTATTTATCAAGTTGATAAGATCTATTAAAGATTA 840
```

QY 841 CCTAGTGATATAAGCATTATTTCAGAGGAAA 873  
 Db 841 CCTAGTGATATAAGCATTATTTCAGAGGAAA 873

RESULT 6  
 US-10-303-162-11  
 ; Sequence 11, Application US/10303162  
 ; Publication No. US20030157658A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wakarchuk, Michel  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/303,162  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 11  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(873)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:4  
 US-10-303-162-11

Query Match 100.0%; Score 873; DB 7; Length 873;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
 Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTCAGG 60  
 Db 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTCAGG 60

QY 61 CTACCAAAATGATTTTGTATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 Db 61 CTACCAAAATGATTTTGTATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTTGGAGCAATACTAC 180  
 Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTTGGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAACTAATATGTTCTAAT 240  
 Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAACTAATATGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTCT 300  
 Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTCT 300

QY 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360  
 Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360

QY 361 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGTCTATATGTCAGTAGCC 420  
 Db 361 TTTACGAAATTTATTTCAATCAAGAAATTAATCTCAGGGTCTATATGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAAATGGGTCA 480  
 Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540  
 Db 481 TCTTATGCTTTTGTATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAT 540

QY 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600  
 Db 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600

QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660  
 Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660

QY 661 ATAGAACTAGCGCCCAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT 720  
 Db 661 ATAGAACTAGCGCCCAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT 720

QY 721 AAGATATATCTCATACCTCTAGTGAGCTTTATGAAAAATTTTCAAAAAATATTAAATTTT 780  
 Db 721 AAGATATATCTCATACCTCTAGTGAGCTTTATGAAAAATTTTCAAAAAATATTAAATTTT 780

QY 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
 Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAAGCATTATTTCAGAGGAAA 873  
 Db 841 CCTAGTGATATAAAGCATTATTTCAGAGGAAA 873

RESULT 7  
 US-10-735-419-11  
 ; Sequence 11, Application US/10735419  
 ; Publication No. US20040180406A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/735,419  
 ; CURRENT FILING DATE: 2003-12-11  
 ; PRIOR APPLICATION NUMBER: US/09/816,028A  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 11  
 ; LENGTH: 873  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(873)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:4  
 US-10-735-419-11

Query Match 100.0%; Score 873; DB 9; Length 873;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
 Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTCAGG 60  
 Db 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTCAGG 60

QY 61 CTACCAAAATGATTTTGTATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
 Db 61 CTACCAAAATGATTTTGTATGCTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120



```
QY 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTGGTTCCTTTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTGGTTCCTTTCTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGGAACCTAATATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGGAACCTAATATGTTCTTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATTTTCT 300
QY 301 GATGCTCAATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360
Db 301 GATGCTCAATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAAA 360
QY 361 TTTCACGAAATTTTATCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
Db 361 TTTCACGAAATTTTATCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAAACAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAAACAGTCTTTTAGCAAAATTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATATCATACAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATATCATACAAGAAAAATAACTACACT 720
QY 721 AAAGATATACATCATACCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTAATTT 780
Db 721 AAAGATATACATCATACCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTAATTT 780
QY 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
QY 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
```

## RESULT 8

```
US-10-820-536-11
; Sequence 11, Application US/10820536
; Publication No. US20040203103A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
```

```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-820-536-11
```

```
Query Match 100.0%; Score 873; DB 9; Length 873;
Best Local Similarity 100.0%; Pred. No. 2,4e-138;
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTCGATTATTCAGG 60
Db 1 ATGAAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTCGATTATTCAGG 60
QY 61 CTACCAATGATTTTGTATGATGTTATAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Db 61 CTACCAATGATTTTGTATGATGTTATAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
QY 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTGGTTCCTTTCTTTGAGCAATACTAC 180
Db 121 CTTGGTAAAAAATGCAAGCAGTGTGTTTACACCCCTGGTTCCTTTCTTTGAGCAATACTAC 180
QY 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGGAACCTAATATGTTCTTAAT 240
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGGAACCTAATATGTTCTTAAT 240
QY 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATTTTCT 300
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATCTTTTACGATATTTTCT 300
QY 301 GATGCTCAATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTTAAATCTTTTAAAA 360
Db 301 GATGCTCAATTTGGGATGATGTTTTTAAACAACTTAAAGAAATTTTAAATCTTTTAAAA 360
QY 361 TTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
Db 361 TTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGAGTAGCC 420
QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAATGGGTCA 480
QY 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
QY 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
Db 541 GATCGCTCACACTATATCGGCATAGTAAATAACAGATATATAAGCTTTTGAATTTCTA 600
QY 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAAACAGTCTTTTAGCAAAATTT 660
Db 601 GAAAAAATTTACAAAAATAAACTATATGCTTATGCTCTTAAACAGTCTTTTAGCAAAATTT 660
QY 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATATCATACAAGAAAAATAACTACACT 720
Db 661 ATAGAACTAGCGCCAAATTTAAATTTAAATTTTATATCATACAAGAAAAATAACTACACT 720
QY 721 AAAGATATACATCATACCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTAATTT 780
Db 721 AAAGATATACATCATACCTTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATTTAATTT 780
QY 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Db 781 AAAAAATATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840
```

QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAA 873  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAA 873

RESULT 9  
US-10-845-408-11  
; Sequence 11, Application US/10845408  
; Publication No. US20040203112A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/845,408  
; PRIOR FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-845-408-11

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTCTGGAATCGAACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTTCTGGAATCGAACCAAGTTTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120  
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120

QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATACTAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATACTAC 180

QY 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCCAAATCAAGAATATGAGCCGAACTAATTTATGTTCTAAT 240

QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTATTTCT 300

QY 301 GATGCTATTTGGGATGATTTTTTTAAACAACTTAAAGATTTAAATGCTTTATTTAAA 360  
Db 301 GATGCTATTTGGGATGATTTTTTTAAACAACTTAAAGATTTAAATGCTTTATTTAAA 360

QY 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
Db 361 TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTTTCAAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATGATTTTTTCAAAAATGGGTCA 480

QY 481 TCTTATGCTTTTGATATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540  
Db 481 TCTTATGCTTTTGATATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

QY 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTAGAATTTCTA 600  
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTAGAATTTCTA 600

QY 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660  
Db 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660

QY 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAAATACTACT 720  
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAAATACTACT 720

QY 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTTT 780  
Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTTT 780

QY 781 AAAAAAATAAAAAATAAAGAAATGTTTATTAACAAGTTGATAAAGAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAAAAATAAAGAAATGTTTATTAACAAGTTGATAAAGAGATCTATTAAAGATTA 840

QY 841 CCTAGTGATATAAGCATTATTTCAAAGGAAA 873  
Db 841 CCTAGTGATATAAGCATTATTTCAAAGGAAA 873

## RESULT 10

US-10-845-412-11  
; Sequence 11, Application US/10845412  
; Publication No. US20040203113A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/845,412  
; PRIOR FILING DATE: 2004-05-12  
; PRIOR APPLICATION NUMBER: US/10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-845-412-11

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTTATGCTGGAATGCAACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTTATGCTGGAATGCAACCAAGTTTAAAGAAATGATTATTCAAGG 60

QY 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

```

Db 61 CTACCAAAATGATTTTGGATGATTTAGATGATTAATCAATTTATTTTGAAGATAAATACAT 120
Qy 121 CTGTGTAATAAATGCAAAAGCAGTGTTTACACCCCTGGTTCTTCTTTGAGCAATACAT 180
Db 121 CTGTGTAATAAATGCAAAAGCAGTGTTTACACCCCTGGTTCTTCTTTGAGCAATACAT 180
Qy 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAATATGTTCTTAAT 240
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGAGCCGAACTAATATGTTCTTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGAAGAACTTTTAAAGCTTTTAAAGCTTTTAA 300
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGAAGAACTTTTAAAGCTTTTAAAGCTTTTAA 300
Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAAGCTTTTAAAGCTTTTAA 360
Db 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAAGCTTTTAAAGCTTTTAA 360
Qy 361 TTTACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTACGAAATTTATTTCAATCAAAAGAAATTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTGATACCAAAATGAAATTTTAAAGAACTTTTAAAGCTTTTAAAGCTTTTAA 540
Db 481 TCTTATGCTTTGATACCAAAATGAAATTTTAAAGAACTTTTAAAGCTTTTAAAGCTTTTAA 540
Qy 541 GATCGCTCACACTATATCGGATAGTAAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 600
Db 541 GATCGCTCACACTATATCGGATAGTAAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 600
Qy 601 GAAAAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 660
Db 601 GAAAAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 660
Qy 661 ATAGAACTTAGCGCCAAATTTAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 720
Db 661 ATAGAACTTAGCGCCAAATTTAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 720
Qy 721 AAAGATATACATACATCTTCTAGTGGGCTTTATGAGAAATTTTCAAAAGCTTTTAAAGCTTTT 780
Db 721 AAAGATATACATACATCTTCTAGTGGGCTTTATGAGAAATTTTCAAAAGCTTTTAAAGCTTTT 780
Qy 781 AAAAAATATAAATTAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTT 840
Db 781 AAAAAATATAAATTAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTT 840
Qy 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

```

RESULT 11  
 US-10-846-219-11  
 ; Sequence 11, Application US/10846219  
 ; Publication No. US20040219638A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/846,219  
 ; CURRENT FILING DATE: 2004-05-14  
 ; PRIOR APPLICATION NUMBER: US/09/816,028  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01

```

; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-846-219-11

```

```

Query Match 100.0%; Score 873; DB 9; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.4e-138; Indels 0; Gaps 0;
Matches 873; Conservative 0; Mismatches 0;

Qy 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAAAGAAATTCATTATTCAAGG 60
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAAAGAAATTCATTATTCAAGG 60
Qy 61 CTACCAAAATGATTTTATGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACAT 120
Db 61 CTACCAAAATGATTTTATGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACAT 120
Qy 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCTTCTTTGAGCAATACAT 180
Db 121 CTGTGTAATAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCTTCTTTGAGCAATACAT 180
Qy 181 ACTTTAAACATTTTAAATCAAAATCAAGATATGAGCCGAACTAATATGTTCTTAAT 240
Db 181 ACTTTAAACATTTTAAATCAAAATCAAGATATGAGCCGAACTAATATGTTCTTAAT 240
Qy 241 TACACCAAGCTCATCTAGAAATGAAATTTTGAAGAACTTTTAAAGCTTTTAAAGCTTTT 300
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGAAGAACTTTTAAAGCTTTTAAAGCTTTT 300
Qy 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAAGCTTTTAAAGCTTTT 360
Db 301 GATGCTCATTTGGGATGATTTTAAACAACTTTAAAGAAATTTAAAGCTTTTAAAGCTTTT 360
Qy 361 TTTACGAAATTTATTTCAATCAAAAGAAATTTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Db 361 TTTACGAAATTTATTTCAATCAAAAGAAATTTTACCTCAGGGGTCTATATGTTGCGAGTAGCC 420
Qy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 481 TCTTATGCTTTGATACCAAAATGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 540
Db 481 TCTTATGCTTTGATACCAAAATGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAA 540
Qy 541 GATCGCTCACACTATATCGGATAGTAAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 600
Db 541 GATCGCTCACACTATATCGGATAGTAAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 600
Qy 601 GAAAAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 660
Db 601 GAAAAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 660
Qy 661 ATAGAACTTAGCGCCAAATTTAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 720
Db 661 ATAGAACTTAGCGCCAAATTTAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 720
Qy 721 AAAGATATACATACATCTTCTAGTGGGCTTTATGAGAAATTTTCAAAAGCTTTTAAAGCTTTT 780
Db 721 AAAGATATACATACATCTTCTAGTGGGCTTTATGAGAAATTTTCAAAAGCTTTTAAAGCTTTT 780
Qy 781 AAAAAATATAAATTAAGAAATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTTTAAAGCTTT 840

```

Db 781 AAAAAAATAAATAAAGAAATGTTTATTACAGTGTATTAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATTAAGCAATTATTTCAGGAAAA 873  
Db 841 CCTAGTGATATTAAGCAATTATTTCAGGAAAA 873

RESULT 12

US-10-821-604-11  
; Sequence 11, Application US/10821604  
; Publication No. US20040229263A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/821,604  
; CURRENT FILING DATE: 2004-04-08  
; PRIOR APPLICATION NUMBER: 10/303,128  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-821-604-11

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCAAGG 60  
QY 61 CTACCAAAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
Db 61 CTACCAAAATGATTTTGATGTTATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120  
QY 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATAC 180  
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATAC 180  
QY 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAATTAATGTTGTTCTAAT 240  
Db 181 ACTTTAAACATTTAATCAAAATCAAGATATGACCGAATTAATGTTGTTCTAAT 240  
QY 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATTTTCT 300  
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTTCAGATTTTCT 300  
QY 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360  
Db 301 GATGCTCAATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTTATTTAA 360  
QY 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTCAGTAGCC 420

QY 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 481 TCTTATGCTTTTGCATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540  
Db 481 TCTTATGCTTTTGCATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540  
QY 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTAAATAATTTCTA 600  
Db 541 GATCGCTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTTAAATAATTTCTA 600  
QY 601 GAAAAAATTTACAAATTAATACTATATGCTTATGCTCTCAACAGCTTTTAAAGCAATTTT 660  
Db 601 GAAAAAATTTACAAATTAATACTATATGCTTATGCTCTCAACAGCTTTTAAAGCAATTTT 660  
QY 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATAACTACACT 720  
QY 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780  
Db 721 AAAGATATATCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780  
QY 781 AAAAAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
Db 781 AAAAAAATAAATAAAGAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 841 CCTAGTGATATAAAGCAATTATTTCAGGAAAA 873  
Db 841 CCTAGTGATATAAAGCAATTATTTCAGGAAAA 873

RESULT 13

US-10-847-983-11  
; Sequence 11, Application US/10847983  
; Publication No. US20040229272A1  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/847,983  
; CURRENT FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-847-983-11

Query Match 100.0%; Score 873; DB 9; Length 873;  
Best Local Similarity 100.0%; Pred. No. 2.4e-138;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTTATTTCAAGG 60

Qy	61	CTACCAATGATTTTGGATGTATTTAGATGTAACTAAATTTATTTTGAAGATAAATACTAT	120
Db	61	CTACCAATGATTTTGGATGTATTTAGATGTAACTAAATTTATTTTGAAGATAAATACTAT	120
Qy	121	CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCCTCTTTGAGCAATACTAC	180
Db	121	CTTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCCTCTTTGAGCAATACTAC	180
Qy	181	ACTTTAAAAATTTAAATCCAAAAATCAGAATATGAGCCGAACCTAATATGTGTTCTAAT	240
Db	181	ACTTTAAAAATTTAAATCCAAAAATCAGAATATGAGCCGAACCTAATATGTGTTCTAAT	240
Qy	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTGAAGATTTTAA	300
Db	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTGAAGATTTTCT	300
Qy	301	GATGCTCATTTGGGATATGATTTTTTTTAAAAACAATTTAAAGAAATTTAAATTTTAA	360
Db	301	GATGCTCATTTGGGATATGATTTTTTTTAAAAACAATTTAAAGAAATTTAAATTTTAA	360
Qy	361	TTTTCACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGTCAGTAGCC	420
Db	361	TTTTCACGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGTCTATATGTGTCAGTAGCC	420
Qy	421	ATAGCCCTTAGGATACAAAGAAATTTATCTCTTCGGGAATTTGATTTTTTATCAAAATGGTCA	480
Db	421	ATAGCCCTTAGGATACAAAGAAATTTATCTCTTCGGGAATTTGATTTTTTATCAAAATGGTCA	480
Qy	481	TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
Db	481	TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
Qy	541	GATCGCTCACATATATCGGACATAGTAAAAATCAGATATAAAGCTTTAGAAATTTCTA	600
Db	541	GATCGCTCACATATATCGGACATAGTAAAAATCAGATATAAAGCTTTAGAAATTTCTA	600
Qy	601	GAATAAACTTTACAAAAATAAAACTATATTTGCTTATGTCCTAAACAGCTTTTAGCAAAATTT	660
Db	601	GAATAAACTTTACAAAAATAAAACTATATTTGCTTATGTCCTAAACAGCTTTTAGCAAAATTT	660
Qy	661	ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATACTACACT	720
Db	661	ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATACTACACT	720
Qy	721	AAAGATATCTCATACCTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATATTTAATTTT	780
Db	721	AAAGATATCTCATACCTCTAGTGAGGCTTTATGGAATAATTTTCAAAAAATATTTAATTTT	780
Qy	781	AAAAAATAAAAAATTTAAAGAAAAATGTTTTATTACAAGTTGATAAAAGATCTATTAGATTA	840
Db	781	AAAAAATAAAAAATTTAAAGAAAAATGTTTTATTACAAGTTGATAAAAGATCTATTAGATTA	840
Qy	841	CCTAGTGATATATAAGCATTTATTTCAAAGGAAAA	873
Db	841	CCTAGTGATATATAAGCATTTATTTCAAAGGAAAA	873

## RESULT 14

US-10-821-573-11  
 ; Sequence 11, Application US/10821573  
 ; Publication No. US20040229313A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/821,573  
 ; CURRENT FILING DATE: 2004-04-08  
 ; PRIOR APPLICATION NUMBER: 10/303,128  
 ; PRIOR FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: US/09/816,028

	Qy	Db	Qy	Db
781	AAAAAATAAAAAATTAAGAAAAATCTTTATTATCAAGTTGATATAAAGATCTATTAAGATTA	840		
781	AAAAAATAAAAAATTAAGAAAAATCTTTATTATCAAGTTGATATAAAGATCTATTAAGATTA	840		
841	CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA	873		
841	CCTAGTGATATAAAGCAATTATTTCAAAGGAAAA	873		

RESULT 15

```

US-10-850-807-11
; Sequence 11, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-850-807-11

```

Query Match	100.0%;	Score 873;	DB 9;	Length 873;
Best Local Similarity	100.0%;	Pred. No. 2.4e-138;		
Matches 873;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAAAACGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60	
DB	1	ATGAAAACGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG	60	
QY	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTAAGATAAAATCTAT	120	
DB	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTGTAAGATAAAATCTAT	120	
QY	121	CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATACTAC	180	
DB	121	CTTGGTAAAAATGCAAAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATACTAC	180	
QY	181	ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTTAAT	240	
DB	181	ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTTCTTAAT	240	
QY	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATATTATTCCT	300	
DB	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATATTATTCCT	300	
QY	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360	
DB	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA	360	
QY	361	TTTTACGAAATTTATTTCATCAAGAAATTAACCTCAGGGTCTATATATGTGCGATGACC	420	
DB	361	TTTTACGAAATTTATTTCATCAAGAAATTAACCTCAGGGTCTATATATGTGCGATGACC	420	

Search completed: July 19, 2006, 15:59:50  
Job time : 998.192 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2006, 15:26:04 ; Search time 141.172 Seconds

(without alignments)  
8883.510 Million cell updates/sec

Title: US-10-734-719-11

Perfect score: 873

Sequence: 1 atgaaaaaagtattattgc.....agcattatttcaaggagaaaa 873

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 899801 seqs, 718270062 residues

Total number of hits satisfying chosen parameters: 1799602

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 2: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 3: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 4: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 5: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 6: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 7: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 8: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*
- 9: /EMC\_Celerra\_SIDS3/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	70.2	8.0	13286	6	US-10-517-441-746
C 2	67	7.7	13286	6	US-10-517-441-472
C 3	67	7.7	13286	6	US-10-517-441-688
C 4	65.6	7.5	5286	6	US-10-517-441-294
C 5	65.6	7.5	5286	6	US-10-517-441-568
C 6	65.6	7.5	8900	6	US-10-517-441-428
C 7	65.6	7.5	8900	6	US-10-517-441-702
C 8	63.8	7.3	10224	6	US-10-517-441-638
C 9	63.8	7.3	11001	6	US-10-517-441-779
C 10	63	7.2	5286	6	US-10-517-441-294
C 11	63	7.2	5286	6	US-10-517-441-568
C 12	62.2	7.1	19634	6	US-10-517-441-414
C 13	61	7.0	6001	6	US-10-517-441-771
C 14	60.8	7.0	2501	6	US-10-517-441-562
C 15	60.6	6.9	2522	7	US-11-218-305-24510
C 16	60.6	6.9	10224	6	US-10-517-441-364
C 17	60.6	6.9	11001	6	US-10-517-441-505
C 18	60	6.9	8093	6	US-10-517-441-308
C 19	60	6.9	8169	6	US-10-517-441-269
C 20	59.4	6.8	9859	6	US-10-517-441-455
C 21	59.4	6.8	9859	6	US-10-517-441-729
C 22	59	6.8	8093	6	US-10-517-441-307
C 23	58.8	6.7	12610	6	US-10-517-441-442
C 24	58.4	6.7	5493	6	US-10-517-441-691

C 25	58.4	6.7	8093	6	US-10-517-441-582	Sequence 582, App
C 26	58.4	6.7	8169	6	US-10-517-441-543	Sequence 543, App
C 27	58.2	6.7	5286	6	US-10-517-441-293	Sequence 293, App
C 28	58.2	6.7	5286	6	US-10-517-441-567	Sequence 567, App
C 29	57.8	6.6	6001	6	US-10-517-441-497	Sequence 497, App
C 30	57.6	6.6	2501	6	US-10-517-441-288	Sequence 288, App
C 31	57.6	6.6	8093	6	US-10-517-441-581	Sequence 581, App
C 32	57.6	6.6	12610	6	US-10-517-441-441	Sequence 441, App
C 33	57.4	6.6	1000	8	US-11-266-748A-281383	Sequence 281383,
C 34	57.4	6.6	1000	8	US-11-266-748A-308023	Sequence 308023,
C 35	57.4	6.6	16579	6	US-10-517-441-563	Sequence 563, App
C 36	57.2	6.6	9859	6	US-10-517-441-456	Sequence 456, App
C 37	57.2	6.6	14147	6	US-10-517-441-539	Sequence 539, App
C 38	56.6	6.5	5286	6	US-10-517-441-24	Sequence 24, Appl
C 39	56.6	6.5	8020	6	US-10-517-441-633	Sequence 633, App
C 40	56.6	6.5	10865	6	US-10-517-441-738	Sequence 738, App
C 41	56.4	6.5	4022	6	US-10-517-441-732	Sequence 732, App
C 42	56.4	6.5	8172	6	US-10-517-441-443	Sequence 443, App
C 43	56.4	6.5	8172	6	US-10-517-441-717	Sequence 717, App
C 44	56.2	6.4	17897	6	US-10-517-441-344	Sequence 344, App
C 45	56.2	6.4	17897	6	US-10-517-441-618	Sequence 618, App

#### ALIGNMENTS

##### RESULT 1

US-10-517-441-746/c

; Sequence 746, Application US/10517441

; Publication No. US20060121467A1

; GENERAL INFORMATION:

; APPLICANT: FOEKENS, John

; APPLICANT: HARBECK, Nadia

; APPLICANT: KOENIG, Thomas

; APPLICANT: MAIER, Sabine

; APPLICANT: MARTENS, John

; APPLICANT: MODEL, Fabian

; APPLICANT: NIMMICH, Inko

; APPLICANT: RUJAN, Tamas

; APPLICANT: SCHMITT, Armin

; APPLICANT: SCHMITT, Manfred

; APPLICANT: LOOK, Maxime P.

; APPLICANT: MARX, Almuth

; APPLICANT: HOEFER, Heinz

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cancer

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE: 47675-93

; CURRENT APPLICATION NUMBER: US/10/517,441

; PRIOR FILING DATE: 2004-12-11

; PRIOR APPLICATION NUMBER: PCT/EP2003/010881

; PRIOR FILING DATE: 2003-10-01

; PRIOR APPLICATION NUMBER: DE 10317955.0

; PRIOR FILING DATE: 2003-04-17

; PRIOR APPLICATION NUMBER: DE 10300096.8

; PRIOR FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: DE 10245779.4

; PRIOR FILING DATE: 2002-10-01

; NUMBER OF SEQ ID NOS: 2147

; SEQ ID NO 746

; LENGTH: 13286

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-517-441-746

Query Match 8.0%; Score 70.2; DB 6; Length 13286;

Best Local Similarity 47.1%; Pred. No. 0.00089;

Matches 317; Conservative 0; Mismatches 348; Indels 8; Gaps 3;

QY 175 TACTACACTTTAAACATTTCCTAAATCAAGAAATATGAGACCGAATATATGTGT 234

DB 10482 TCCTTTATTACTAAAAAAATCCAACTACCTCTAAACCTACATATTTCTT 10423





```

; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxine P.
; APPLICANT: MARX, Almuth
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR FILING DATE: 2003-04-17
; PRIOR FILING DATE: 2003-01-07
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 688
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-688

Query Match
Best Local Similarity 7.7%; Score 67; DB 6; Length 19634;
Matches 326; Conservative 0; Mismatches 370; Indels 7; Gaps 3;

QY 159 TTCTCTTTGAGCAATACACTTTTAAACATTTAAATCCAAATATGAGATATGAGAC 218
DB 7370 TTACTTAAACAAAAACACACATAAATAATATCTTAAATAATTTTATATAAAT 7311
QY 219 CGAATAATATGTTCTTAATTAACAACAGCTCATCTAGAAATGAAATTTGTAA 278
DB 7310 ATAAACCAACCTAAATAATCAATTAACAAAACTTAAATTAATTTTCTATTTATA 7251
QY 279 AACTTTTACGATATTTTCTGATGCTCATTTGGGATATGATTTTAAACACTAA 338
DB 7250 AACCAATATTTCTAATCATTCTATCTTAAACATATAAATTCATTAATAAAA 7191
QY 339 AGAATTTAATGCTTATTTTAAATTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGG 398
DB 7190 AAAATTTTAAACACTATTTTCATTTTCTTTCCCATCAATATATCACTAAATTT 7131
QY 399 GGTCTATATGTTGAGTAGGATAGCCCTAGATACAAAGAAATTTATCTTTGGGAT 458
DB 7130 ATACCAAAAAATTTTACAAAAACACAAATTTACTTTATATAATTTATTTATAAAC 7075
QY 459 TGATTTTATCAAAATGGTCTATCTTATGCTTTTATGATACCAACCAAGAAATCTTTAA 518
DB 7074 TCTATATATCTACTT--CACATCTTACTCTCAATTTTAACTTTTAACTTATCTACA 7017
QY 519 ACTAGCCCTGATTTTAAATAATGATCGCTCACACTATATC--GGACATAGTAAATAACAG 577
DB 7016 ATTAACCTTAAATAATCAACAATACTAAACAAATCTTATAAATTTACTAATTTTC 6957
QY 578 ATATAAAGCTTTAGAAATTTCTAGAAAAAATTTACAAATATAAATTTATGCTTTATGTC 637
DB 6956 ATAAACACCCCAAACTCTTCTTCAATCATTTTAAATAACAAAAATTTAATATAC 6897
QY 638 CTACAGCTTTTAGCAATTTTATAGCACTAGCCCAATTTAAATCAAAATTTTATCA 697
DB 6896 ACAAATAAATAATAAACCCTTCTCAATCTTAAATCTTAAATTTAAATTTAAAT 6837

```

```

QY 698 TACAAGAAAAAATAACTACTAAAGATATATCTCATACCTTCTAGTGAGGCTTATGAA 757
DB 6836 TTAACAATAATAAAAAATAAAAAATCTAATATATCTTAACTTATAAACAATCTTTTACC 6777
QY 758 AATTTTCAAAAAATATTAATTTTAAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGT 817
DB 6776 TTTTATTAATAATTTTACCTTTAATAATTTCTTCTATTTATTTATTTTAAAAAAA 6717
QY 818 TGATAAAGATCTATTAAAGATTACCTAGTGTATATAAAGCAATTA 860
DB 6716 ACCTAAAAAACATAAAACCTTAAATTTACTAAAAAAAACAATA 6674

RESULT 4
US-10-517-441-294/c
; Sequence 294, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxine P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFLE, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR FILING DATE: 2003-04-17
; PRIOR FILING DATE: 2003-01-07
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 294
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-294

Query Match
Best Local Similarity 7.5%; Score 65.6; DB 6; Length 5286;
Matches 364; Conservative 0; Mismatches 454; Indels 6; Gaps 2;

QY 40 AAAGAAATGATTTTCAAGGCTTACCAATGATTTTGCATGTTTGTAGATGATTAATTT 99
DB 5175 ATATAAATAATATATATAATATATATATATATATATATATATATATATATATATA 5116
QY 100 TATTTTGAAGATAAATACTATCTTGTAAAAAATGCAACAGTGTTTTACACCCCTGGT 159
DB 5115 TATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 5056
QY 160 TTCTCTTTGAGCAATACACTTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACC 219
DB 5055 ATAATATATATATATATATATATATATATATATATATATATATATATATATATATA 4996
QY 220 GAACATATATGTTTCTAATTTACAAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAA 279
DB 4995 TATTATATATATATATATATATATATATATATATATATATATATATATATATATA 4936

```

```
QY 280 ACTTTTACGATTATTTTCCTGATGCTCATTTGGGATATGATTTTTTTAAACAACCTTAA 339
Db 4935 ATATATTATATATATATATATATATATATATATATATATATATATATATATATATAT 4876

QY 340 GAAATTAATGCTTATTTTAAATTTTCCAGAAATTTTAAATTTTCAATCAAGAATTTACCTCAGGG 399
Db 4875 AATATTAAATATATATATATATATATATATATATATATATATATATATATATATATAT 4816

QY 400 GTCTATATGTGTCAGTACCATGCTTGTAGATACAAAGAAATTTATCTTTTCGGGAATTT 459
Db 4815 AACAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4761

QY 460 GATTTTATCAAAATGGGCTATCTTATGCTTTTGTATACCAACAGAAATCTTTTAA 519
Db 4760 ATTATATAATATATATATATATATATATATATATATATATATATATATATATATATAT 4701

QY 520 CTAGCCCTGATTTTAAATAATGATCGCTCACACTATATCGGCATATAGTAAATAACAGAT 579
Db 4700 ATATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4641

QY 580 ATAAAGCTTTAGAAATTTCTAGAAAAAATTTTACAAAAAATTTTAACTATATTTGCTTATGCTCT 639
Db 4640 TATATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4581

QY 640 AACAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTTAAATTTTCAATTTTATCATTA 699
Db 4580 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4522

QY 700 CAAGAAAAATAAATACTACACTAAAGATATATCTCATACCTTCTAGTGGGCTTATGAAAA 759
Db 4521 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4462

QY 760 TTTTCAAAAAATTTTAAATTTTAAAAAATAAAAAATTTTAAAGAAAAATTTTATTTACAAAGTTG 819
Db 4461 TATAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4402

QY 820 ATAAAGATCTTTAAGATTACCTAGTGNATATAAGCATTTATTT 863
Db 4401 ATATATAATATATATATATATATATATATATATATATATATATATATATATATATATAT 4358

RESULT 5
US-10-517-441-568/c
; Sequence 568, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEBKENS, John
; APPLICANT: HABBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMWICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOFELER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010861
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
```

```
; SEQ ID NO 568
; LENGTH: 5286
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-568
```

```
Query Match 7.5%; Score 65.6; DB 6; Length 5286;
Best Local Similarity 44.2%; Pred. No. 0.0051;
Matches 364; Conservative 0; Mismatches 454; Indels 6; Gaps 2;
```

```
QY 40 AAAGAAATTCATTTTCAAGGCTACCAATGATTTTGTATGATTTTGTAGATGATTAATCAATTTT 99
Db 5175 ATATAAAAATATATATATATATATATATATATATATATATATATATATATATATATATATTA 5116

QY 100 TATTTTGAAGATAAATACTATCTTGTGTAATAAATGCAAGAGCAGTGTTTTACACCCCTGGT 159
Db 5115 TATAATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5056

QY 160 TTCTTTCTTTGAGCAATACCTACACTTTTAAACATTTTAAATCCAAATCAAGAATATGAGACC 219
Db 5055 ATAAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4996

QY 220 GAACTAATATATGTTCTTAATTTACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAA 279
Db 4995 TATTATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4936

QY 280 ACTTTTACGATTTATTTTCTGATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAA 339
Db 4935 ATATATTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4876

QY 340 GAAATTAATGCTTTTAAATTTTCAACGAAATTTTCAATCCAAAGAAATTTACCTCAGGG 399
Db 4875 AATATTAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4816

QY 400 GTCTATATGTGTCAGTACCATGAGCTAGAGATCAAAAGAAATTTATCTTTTCGGGAATTT 459
Db 4815 AACCAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4761

QY 460 GATTTTATCAAAATGGGCTATCTTATGCTTTTGTATACCAACAAAGAAATCTTTTAA 519
Db 4760 ATTATATAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4701

QY 520 CTAGCCCTGATTTTAAATAATGATCGCTCACACTATATCGGCATATAGTAAAAATACAGAT 579
Db 4700 ATATTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4641

QY 580 ATAAAGCTTTAGAAATTTCTAGAAAAAATTTTACAAAAAATTTTAACTATATTTGCTTATGCTCT 639
Db 4640 TATATAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4581

QY 640 AACAGTCTTTTAGCAAAATTTTATAGAACTAGCGCCAAATTTTAAATTTTCAATTTTATCATTA 699
Db 4580 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4522

QY 700 CAAGAAAAATAAATACTACACTAAAGATATATCTCATACCTTCTAGTGGGCTTATGAAAA 759
Db 4521 TAAATTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4462

QY 760 TTTTCAAAAAATTTTAAATTTTAAAAAATAAAAAATTTTAAAGAAAAATTTTATTTACAAAGTTG 819
Db 4461 TATAATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4402

QY 820 ATAAAGATCTTTAAGATTACCTAGTGNATATAAGCATTTATTT 863
Db 4401 ATATATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4358
```

```
RESULT 6
US-10-517-441-428/c
; Sequence 428, Application US/10517441
; Publication No. US20060121467A1
```

; GENERAL INFORMATION:  
 ; APPLICANT: FOEKENS, John  
 ; APPLICANT: HARBECK, Nadia  
 ; APPLICANT: KOENIG, Thomas  
 ; APPLICANT: MAIER, Sabine  
 ; APPLICANT: MARTENS, John  
 ; APPLICANT: MODEL, Fabian  
 ; APPLICANT: NIMMICH, Inko  
 ; APPLICANT: RUJAN, Tamas  
 ; APPLICANT: SCHMITT, Armin  
 ; APPLICANT: SCHMITT, Manfred  
 ; APPLICANT: LOOK, Maxime P.  
 ; APPLICANT: MARX, Almuth  
 ; APPLICANT: HOEFER, Heinz  
 ; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
 ; FILE OF INVENTION: proliferative disorders  
 ; FILE REFERENCE: 47675-93  
 ; CURRENT APPLICATION NUMBER: US/10/517,441  
 ; CURRENT FILING DATE: 2004-12-11  
 ; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: DE 10317955.0  
 ; PRIOR FILING DATE: 2003-04-17  
 ; PRIOR APPLICATION NUMBER: DE 10300096.8  
 ; PRIOR FILING DATE: 2003-01-07  
 ; PRIOR APPLICATION NUMBER: DE 10245779.4  
 ; PRIOR FILING DATE: 2002-10-01  
 ; NUMBER OF SEQ ID NOS: 2147  
 ; SEQ ID NO 428  
 ; LENGTH: 8900  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-517-441-428

Query Match 7.5%; Score 65.6; DB 6; Length 8900;  
 Best Local Similarity 50.3%; Pred. No. 0.0055;  
 Matches 187; Conservative 0; Mismatches 184; Indels 1; Gaps 1;  
 QY 503 AAGAAATCTTTTAAACTAGCCCTGATTTTAAATAATGATCGCTCACACTATATCGGAC 562  
 DB 536 AATATATTTTATTTTAACTTCTTAATTAATATATATATATATATATATATATATAT 477  
 QY 563 ATAGTAAAAATACAGATATAAAGCTTTTAAATAATGATCGCTCACACTATATCGGAC 622  
 DB 476 TTACATTTCTTAACTAATAATAATAATAATAATAATAATAATAATAATAATAATA 417  
 QY 623 TATATGCTTATGTC-CTAACAGTCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAA 681  
 DB 416 TATATATTTTATATCTCATATATAAATTTTATATATATTTTCTTAAATATATCTAATTTA 357  
 QY 682 AATTCAAATTTTATCATATACAGAAATAAATACTACATAAGATATATCTACCTCTT 741  
 DB 356 AATTCATATTTTACAAATAAATACTTAAATACTTCTTAAATAATAAATAATAATA 297  
 QY 742 AGTGAGGCTTATGGAATTTTCAAAAAATTTTAAATAATAATAATAATAATAATA 801  
 DB 296 AATTTAAATTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATA 237  
 QY 802 AATGTTTATTAACAGTTGATATAAGATCTTATTAAGATTAACCTAGATATAAGCATTTAT 861  
 DB 236 TACAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 177  
 QY 862 TTCAAAGGAAAA 873  
 DB 176 TTTTAAAAAAA 165

RESULT 7  
 US-10-517-441-702/c  
 ; Sequence 702, Application US/10517441  
 ; Publication No. US20060121467A1

; GENERAL INFORMATION:  
 ; APPLICANT: FOEKENS, John  
 ; APPLICANT: HARBECK, Nadia  
 ; APPLICANT: KOENIG, Thomas  
 ; APPLICANT: MAIER, Sabine  
 ; APPLICANT: MARTENS, John  
 ; APPLICANT: MODEL, Fabian  
 ; APPLICANT: NIMMICH, Inko  
 ; APPLICANT: RUJAN, Tamas  
 ; APPLICANT: SCHMITT, Armin  
 ; APPLICANT: SCHMITT, Manfred  
 ; APPLICANT: LOOK, Maxime P.  
 ; APPLICANT: MARX, Almuth  
 ; APPLICANT: HOEFER, Heinz  
 ; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell  
 ; FILE OF INVENTION: proliferative disorders  
 ; FILE REFERENCE: 47675-93  
 ; CURRENT APPLICATION NUMBER: US/10/517,441  
 ; CURRENT FILING DATE: 2004-12-11  
 ; PRIOR APPLICATION NUMBER: PCT/EP2003/010881  
 ; PRIOR FILING DATE: 2003-10-01  
 ; PRIOR APPLICATION NUMBER: DE 10317955.0  
 ; PRIOR FILING DATE: 2003-04-17  
 ; PRIOR APPLICATION NUMBER: DE 10300096.8  
 ; PRIOR FILING DATE: 2003-01-07  
 ; PRIOR APPLICATION NUMBER: DE 10245779.4  
 ; PRIOR FILING DATE: 2002-10-01  
 ; NUMBER OF SEQ ID NOS: 2147  
 ; SEQ ID NO 702  
 ; LENGTH: 8900  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 US-10-517-441-702

Query Match 7.5%; Score 65.6; DB 6; Length 8900;  
 Best Local Similarity 50.3%; Pred. No. 0.0055;  
 Matches 187; Conservative 0; Mismatches 184; Indels 1; Gaps 1;  
 QY 503 AAGAAATCTTTTAAACTAGCCCTGATTTTAAATAATGATCGCTCACACTATATCGGAC 562  
 DB 536 AATATATTTTATTTTAACTTCTTAATTAATATATATATATATATATATATATAT 477  
 QY 563 ATAGTAAAAATACAGATATAAAGCTTTTAAATAATGATCGCTCACACTATATCGGAC 622  
 DB 476 TTACATTTCTTAACTAATAATAATAATAATAATAATAATAATAATAATAATAATA 417  
 QY 623 TATATGCTTATGTC-CTAACAGTCTTTTAAAGAAATTTTAAAGAAATTTTAAAGAA 681  
 DB 416 TATATATTTTATATCTCATATATAAATTTTATATATATTTTCTTAAATATATCTAATTTA 357  
 QY 682 AATTCAAATTTTATCATATACAGAAATAAATACTACATAAGATATATCTACCTCTT 741  
 DB 356 AATTCATATTTTACAAATAAATACTTAAATACTTCTTAAATAATAAATAATAATA 297  
 QY 742 AGTGAGGCTTATGGAATTTTCAAAAAATTTTAAATAATAATAATAATAATAATA 801  
 DB 296 AATTTAAATTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATA 237  
 QY 802 AATGTTTATTAACAGTTGATATAAGATCTTATTAAGATTAACCTAGATATAAGCATTTAT 861  
 DB 236 TACAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 177  
 QY 862 TTCAAAGGAAAA 873  
 DB 176 TTTTAAAAAAA 165

RESULT 8  
 US-10-517-441-638/c  
 ; Sequence 638, Application US/10517441  
 ; Publication No. US20060121467A1

```
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 638
; LENGTH: 10224
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-638

Query Match      7.3%; Score 63.8; DB 6; Length 10224;
Best Local Similarity 45.2%; Pred. No. 0.012;
Matches 312; Conservative 0; Mismatches 377; Indels 2; Gaps 2;

QY 182 CTTTAAACATTTTAAATCCAAATCAAGAAATGAGACCGAAGTATGTTCTTAAT 241
DB 7181 CTCAAAAAATAAATAAACATAAAAAATAAAAAATAAATTTAAACATATTTATC 7122
QY 242 ACAACCAAGCTCATCTAGAAATGAAATTTGTGAAAACTTTTACGATTATTTTCCTG 301
DB 7121 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7062
QY 302 ATGCTCATTTGGGATATGATTTTTTAAACAACTTTAAAGAAATTTAATGCTTTTAAAT 361
DB 7061 AACCTAATAACCATTTTACTACAAAAAATAAATAAATAAATAAATAAATAAATAA 7002
QY 362 TTCAGCAATTTATTTCAATCAAGAAATTTACTAGGGGTCTATATGTGTCAGTAGCA 421
DB 7001 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6942
QY 422 TAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTTTATCAAAATGGTCA 481
DB 6941 AATACTAACAAAAATCTAATCTAAAAATTTAAATCAATATTTCAACCAATAAATCTAA 6882
QY 482 CTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTCGATTTTAAAGATG 541
DB 6881 AACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6822
QY 542 ATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGATTTCTAG 601
DB 6821 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6763
QY 602 AAAAAACTTACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 661
DB 6762 AAAAAACTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6704
QY 662 TAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAAATAAATAA 721
```

```
DB 6703 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6644
QY 722 AAGATATATCTCATCTCTAGTGGCTTATGAAAAATTTTCAGAAAAATTTTCAAAATAATTTA 781
DB 6643 AACCTATCCCAACAATTTCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 6584
QY 782 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 841
DB 6583 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6524
QY 842 CTAGTGATATAAAGCATTTTCAAGGAAA 872
DB 6523 AAAAAATAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 6493

RESULT 9
US-10-517-441-779/c
; Sequence 779, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
; APPLICANT: FOEKENS, John
; APPLICANT: HARBECK, Nadia
; APPLICANT: KOENIG, Thomas
; APPLICANT: MAIER, Sabine
; APPLICANT: MARTENS, John
; APPLICANT: MODEL, Fabian
; APPLICANT: NIMMICH, Inko
; APPLICANT: RUJAN, Tamas
; APPLICANT: SCHMITT, Armin
; APPLICANT: SCHMITT, Manfred
; APPLICANT: LOOK, Maxime P.
; APPLICANT: MARX, Almuth
; APPLICANT: HOEFER, Heinz
; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE: 47675-93
; CURRENT APPLICATION NUMBER: US/10/517,441
; CURRENT FILING DATE: 2004-12-11
; PRIOR APPLICATION NUMBER: PCT/EP2003/010881
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: DE 10317955.0
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 779
; LENGTH: 11001
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-517-441-779

Query Match      7.3%; Score 63.8; DB 6; Length 11001;
Best Local Similarity 45.2%; Pred. No. 0.012;
Matches 312; Conservative 0; Mismatches 377; Indels 2; Gaps 2;

QY 182 CTTTAAAAATTTAATCCAAATCAAGAAATGAGACCGAAGTATGTTCTTAAT 241
DB 8398 CTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8339
QY 242 ACAACCAAGCTCATCTAGAAATGAAATTTGTGAAAACTTTTACGATTATTTTCCTG 301
DB 8338 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8279
QY 302 ATGCTCATTTGGGATATGATTTTTTAAACAACTTTTAAAGAAATTTAATGCTTTTAAAT 361
DB 8278 AACCTAATAACCATTTTACTACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 8219
QY 362 TTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCCA 421
```

Db 8218 AAAAAAAAAATTAATTTTAAACCATATTCCTTAAACATCTTTTAAACATCCTTAATAA 8159

Qy 422 TAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCATAAATGGGTGAT 481

Db 8158 AATACTAACACAAAATCTAATCTAAATTTTAATCAATATTTTCAACCAATAATCTAA 8099

Qy 482 CTTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAATG 541

Db 8098 AACAAATAAATATATAACCAACCAACCAATAATCAAACTAAACCCAACTTCATCATAA 8039

Qy 542 ATCCCTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTCCTAG 601

Db 8038 AATATAAAATTTAAATCAAA-ATCTACAAAAAACTAAACAAAAAATACAAAAATAAAA 7980

Qy 602 AAAAAAATTTACAAAATAAAATATATTTGCTTATGCTTAAACAGTCTTTTACCAATTTTA 661

Db 7979 AAAAAATTAATAAATAATATACCCCTAAATCAAAAAACAAAC-TTCAATTAATCTA 7921

Qy 662 TAGAACTAGCGCCAAATTTTAAATTTCAAATTTTATCATACAAGAAAAATAAATACACTA 721

Db 7920 AAAAAATAAAAAATAAAAAATAAACACTAAATTCATTTAATAATACAAAAATCACTAAA 7861

Qy 722 AAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAAATTTTA 781

Db 7860 AACCTATCCCAACAACTTTTCTAAAAAAATAAAACCACTAACTAAAAATAAAAAA 7801

Qy 782 AAAAAATAAAAAATTAAGAAATGTTTATTACAGTTGATAAAGATCTTATTAAGATTAC 841

Db 7800 AAAAAATAAAAAATAAAAAATTAATAATATAAATAAATAAATAAATAAATAAATAA 7741

Qy 842 CTAGTGATATAAAGCATTTTTCAAAGGAAA 872

Db 7740 AACAAATAAACCAACCAACCAACATAAA 7710

RESULT 10

US-10-517-441-294

Sequence 294, Application US/10517441

Publication No. US20060121467A1

GENERAL INFORMATION:

APPLICANT: FOSKENS, John

APPLICANT: HARBECK, Nadia

APPLICANT: KOENIG, Thomas

APPLICANT: MAIER, Sabine

APPLICANT: MARTENS, John

APPLICANT: MODEL, Fabian

APPLICANT: NIMMICH, Inko

APPLICANT: RUJAN, Tamas

APPLICANT: SCHMITT, Armin

APPLICANT: SCHMITT, Manfred

APPLICANT: LOOK, Maxime P.

APPLICANT: MARX, Almuth

APPLICANT: HOEFER, Heinz

TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell

FILE REFERENCE: 47675-93

CURRENT APPLICATION NUMBER: US/10/517,441

CURRENT FILING DATE: 2004-12-11

PRIOR APPLICATION NUMBER: PCT/EP2003/010881

PRIOR FILING DATE: 2003-10-01

PRIOR APPLICATION NUMBER: DE 10317955.0

PRIOR FILING DATE: 2003-04-17

PRIOR APPLICATION NUMBER: DE 10300096.8

PRIOR FILING DATE: 2003-01-07

PRIOR APPLICATION NUMBER: DE 10245779.4

PRIOR FILING DATE: 2002-10-01

NUMBER OF SEQ ID NOS: 2147

SEQ ID NO 294

LENGTH: 5286

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:



Db 7130 ATACCGAAAAAATT---TTACGAAAAACACAAATTTACTTTTATATATTTATATAAAC 7075

QY 459 TGATTTTATCAAAATGGGTGATCTTTATGCTTTGATACCAACAGAAATCTTTTAA 518

Db 7074 TCTATATTATCTACTT--CACGTCTTACTCTCATTTTAACTCTTTTAACTTATCTACA 7017

QY 519 ACTAGCCCTGATTTTAAAAATGATCGCTCACACTATATC--GCACATAGTAAATAACAG 577

Db 7016 ATTAAGTCTTAAATAATACACAAATACAAATACTCTATATAATTTACTAAATTTTC 6957

QY 578 ATATAAAGCTTTAGAAATTTTACGAAAAAATTTACAAATAAATTTATATTTGCTTATGTC 637

Db 6956 ATAAACACCCATAAATCTTTCTTCTATATCAATTTTAAATAAACAATAATTTTAAATATAC 6897

QY 638 CTAACAGTCTTTAGCAAAATTTTATAGAACTAGCGCAAAATTTAAATTCCAAATTTATCA 697

Db 6896 ACAAAATRACTATAATAACCCAAATTCCTCAAAATCCTAATAATTTACTTTTAAATTTAAAC 6837

QY 698 TACAAGAAAAAATRACTACACTAAAGATATATCTATACCTTCTAGTGAGGCTTATGGAA 757

Db 6836 TTAACAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6777

QY 758 AATTTTCAAAAAATATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 817

Db 6776 TTTTATAATATATTTTACCTTAATAATTTCTTCTATATATTTTATATATTTTAAATAA 6717

QY 818 TGATAAAGATCTATTAAAGATCTACCTAGTGATATAAAGCATTA 860

Db 6716 ACTTAAAAAACATAAACTTAAATTTACTTAAAAAACGATA 6674

RESULT 13

US-10-517-441-771/c

; Sequence 771, Application US/10517441

; Publication No. US20060121467A1

; GENERAL INFORMATION:

; APPLICANT: FOKENS, John

; APPLICANT: HARBECK, Nadia

; APPLICANT: KOENIG, Thomas

; APPLICANT: MAIER, Sabine

; APPLICANT: MARTENS, John

; APPLICANT: MODEL, Fabian

; APPLICANT: NIMMICH, Inko

; APPLICANT: RUJAN, Tamas

; APPLICANT: SCHMITT, Manfred

; APPLICANT: SCHMITT, Armin

; APPLICANT: LOOK, Maxime P.

; APPLICANT: MARX, Almuth

; APPLICANT: HOEFER, Heinz

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE: 47675-93

; CURRENT APPLICATION NUMBER: US/10517441

; PRIOR FILING DATE: 2004-12-11

; PRIOR APPLICATION NUMBER: PCT/EP2003/010881

; PRIOR FILING DATE: 2003-10-01

; PRIOR APPLICATION NUMBER: DE 10317955.0

; PRIOR FILING DATE: 2003-04-17

; PRIOR APPLICATION NUMBER: DE 10300096.8

; PRIOR FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: DE 10245779.4

; PRIOR FILING DATE: 2002-10-01

; NUMBER OF SEQ ID NOS: 2147

; SEQ ID NO 771

; LENGTH: 6001

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-517-441-771

Query Match 7.0%; Score 61; DB 6; Length 6001;

Best Local Similarity 48.2%; Pred. No. 0.034;

Matches 172; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 517 AAATCTAGCCCTGATTTTAAATAATGATCGCTCACACTATATCGACATAGTAAATAACA 576

Db 3721 AATCAACCCCAAAAAATTTAAATATATTTATCTTAAAAAAATCTTAAATATACATAAAATTTTA 3662

QY 577 GATATAAAGCTTTAGAAATTTTACGAAAAAATTTACAAATAAATTTATATTTGCTTATGT 636

Db 3661 AATAAAAACAATTCCTAAATTTCTTAAAAAAACAAAACTCCTTAAAAAAACA 3602

QY 637 CCTAACAGTCTTTTAGCAAAATTTTATAGAACTAGCGCAAAATTTTAAATTTCAAAATTTTATC 696

Db 3601 CACACCCCATCCCACTCCCAATACCTTAAAAAATCAAAATTAACAAATAATTAACAC 3542

QY 697 ATCAAGAAAAAATAAATACACTACATAAGATATATCTATACCTTCTAGTGAGGCTTATGGA 756

Db 3541 ATTAATAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3482

QY 757 AAATTTTCAAAATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 816

Db 3481 AAAAAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3422

QY 817 TTGATAAAGATCTATTAAAGATTTACCTAGTGATATAAAGCATTTTCAAGGAAAA 873

Db 3421 AAAAAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3365

RESULT 14

US-10-517-441-562/c

; Sequence 562, Application US/10517441

; Publication No. US20060121467A1

; GENERAL INFORMATION:

; APPLICANT: FOKENS, John

; APPLICANT: HARBECK, Nadia

; APPLICANT: KOENIG, Thomas

; APPLICANT: MAIER, Sabine

; APPLICANT: MARTENS, John

; APPLICANT: MODEL, Fabian

; APPLICANT: NIMMICH, Inko

; APPLICANT: RUJAN, Tamas

; APPLICANT: SCHMITT, Armin

; APPLICANT: SCHMITT, Manfred

; APPLICANT: LOOK, Maxime P.

; APPLICANT: MARX, Almuth

; APPLICANT: HOEFER, Heinz

; TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cell

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE: 47675-93

; CURRENT APPLICATION NUMBER: US/10517441

; PRIOR FILING DATE: 2004-12-11

; PRIOR APPLICATION NUMBER: PCT/EP2003/010881

; PRIOR FILING DATE: 2003-10-01

; PRIOR APPLICATION NUMBER: DE 10317955.0

; PRIOR FILING DATE: 2003-04-17

; PRIOR APPLICATION NUMBER: DE 10300096.8

; PRIOR FILING DATE: 2003-01-07

; PRIOR APPLICATION NUMBER: DE 10245779.4

; PRIOR FILING DATE: 2002-10-01

; NUMBER OF SEQ ID NOS: 2147

; SEQ ID NO 562

; LENGTH: 2501

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-517-441-562

Query Match 7.0%; Score 60.8; DB 6; Length 2501;

Best Local Similarity 49.2%; Pred. No. 0.033;

Matches 189; Conservative 0; Mismatches 192; Indels 3; Gaps 1;

QY 492 TGATACCAACAAAGAAATCTTTTAAATAGTCCCTGATTTTAAATATGATCGCTCACA 551





GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:30:08 ; Search time 107.972 Seconds  
(without alignments)  
1232.267 Million cell updates/sec

Title: US-10-734-719-12

Perfect score: 1563

Sequence: 1 MKKVIIAGNGFSLKEIDYR.....KLINDLLRLPSDIKHFKGK 291

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_8.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*
- 9: Geneseq2005s.\*
- 10: Geneseq2006s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1563	100.0	291	6	ABJ18484 Campyloba
2	1557	99.6	291	3	AY97212 Campyloba
3	1557	99.6	291	6	ABJ18482 Campyloba
4	1553	99.4	291	6	ABJ18485 Campyloba
5	1534	98.1	291	3	AY97210 Campyloba
6	1534	98.1	291	6	ABJ18480 Campyloba
7	1524	97.5	291	3	AY97204 Campyloba
8	1524	97.5	291	6	ABJ18479 Campyloba
9	1513	96.8	291	3	AY97211 Campyloba
10	1513	96.8	291	6	ABJ18481 Campyloba
11	804	51.4	294	3	AY97215 Campyloba
12	804	51.4	294	6	ABJ18483 Campyloba
13	749	47.9	430	2	AY45221 Campyloba
14	698.5	44.7	303	6	ABR40195 alpha-2,3
15	119	7.6	958	8	ADN19559 Bacteri
16	112	7.2	2133	8	ADP25432 Plasmodi
17	110	7.0	392	3	ABJ15991 E. coli p
18	109.5	7.0	908	6	ABU19261 Protein e
19	109.5	7.0	1121	3	ABJ18241 Plasmodi
20	109.5	7.0	4134	2	AY31946 Plasmodi
21	109	7.0	1398	3	ABJ18292 Plasmodi
22	107.5	6.9	388	8	ADK16435 Nanoarcha
23	107	6.8	863	6	ADB08930 Alloioocc

24	106.5	6.8	305	6	ABM69267 Photorhab
25	106.5	6.8	1284	6	ABU48910 Protein e
26	106	6.8	1021	2	Aaw98555 H. pylori
27	106	6.8	1815	8	ADP25442 Plasmodi
28	106	6.8	2539	3	AAJ18198 Plasmodi
29	105.5	6.7	400	6	ABU25530 Protein e
30	105.5	6.7	422	4	AAU35736 Helicobac
31	105.5	6.7	422	6	ABR83248 H. pylori
32	105.5	6.7	422	6	ABR83247 H. pylori
33	105.5	6.7	422	6	ABU30795 Protein e
34	104.5	6.7	1411	2	AAJ31948 Plasmodi
35	104.5	6.7	1417	2	AAJ31947 Plasmodi
36	104.5	6.7	1610	7	ADB46133 Plasmodi
37	104	6.7	822	5	ABJ77605 AmpPV sec
38	104	6.7	1817	3	AAJ18255 Plasmodi
39	104	6.7	2013	3	AAJ18265 Plasmodi
40	103.5	6.6	1346	3	AAJ18236 Plasmodi
41	103	6.6	1182	3	AAJ18288 Plasmodi
42	103	6.6	1817	3	AAJ18301 Plasmodi
43	102.5	6.6	1062	8	ADP25429 Plasmodi
44	101.5	6.5	295	7	ADC96625 E. faeciu
45	101	6.5	340	5	ABB54110 Lactococc

## ALIGNMENTS

RESULT 1

ABJ18484

ID ABJ18484 standard; protein; 291 AA.

XX AC ABJ18484;

XX DT 07-FEB-2003 (first entry)

XX DE Campylobacter jejuni bifunctional sialtransferase cstII #5.

XX KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;

KW GalNac transferase; N-Acetylgalactosamine transferase;

KW galactosyltransferase; sialyltransferase; sialic acid synthase;

KW cytidine 5'-monophosphate sialic acid synthetase;

KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;

KW ganglioside mimetics; inflammation; tumour metastasis.

XX OS Campylobacter jejuni.

XX PN WO200274942-A2.

XX PD 26-SEP-2002.

XX PF 22-FEB-2002; 2002WO-CA0000229.

XX PR 21-MAR-2001; 2001US-00816028.

XX XX (CANA ) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX XX WPI; 2003-040554/03.

XX DR N-PSDB; ABT13670.

XX PT New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.

XX PS Disclosure; Page 99; 107pp; English.

XX CC The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an: acyltransferase; glycosyltransferase; GalNAC (N-Acetylgalactosamine) transferase; galactosyltransferase; sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein

CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present amino acid sequence represents a Campylobacter  
CC jejuni protein of the invention  
XX  
SQ Sequence 291 AA;

Query Match 100.0%; Score 1563; DB 6; Length 291;  
Best Local Similarity 100.0%; Pred. No. 2.4e-148;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLKCKKCAVFTPGFFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLKCKKCAVFTPGFFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 120  
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 120  
QY 121 PHEIYFNQIRITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 180  
DB 121 PHEIYFNQIRITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 180  
QY 181 DRSHYIGHSKNTDIALKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIALKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHYPFGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHYPFGK 291

RESULT 2  
ID AAY97212 standard; protein; 291 AA.  
XX  
AC AAY97212;  
XX  
DT 12-SEP-2003 (revised)  
DT 22-DEC-2000 (first entry)  
XX  
DE Campylobacter jejuni O:19 serotype CstII sialyltransferase.  
XX  
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
KW immunity; immunogen; ganglioside.  
XX  
OS Campylobacter jejuni; O:19 serotype.  
XX  
PN WO200046379-A1.  
XX  
PD 10-AUG-2000.  
XX  
PF 01-FEB-2000; 2000WO-CAC000086.  
XX  
PR 01-FEB-1999; 99US-0118213P.  
PR 31-JAN-2000; 2000US-00495406.  
XX  
PA (CANADA) NAT RES COUNCIL CANADA.  
XX  
PI Gilbert M, Wakarchuk WW;  
XX  
DR WPI; 2000-524418/47.  
DR N-PSDB; AAS53726.  
XX  
PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic

PT reagents and as immunogen for producing antibodies.  
XX  
PS Disclosure; Page 96-97; 120pp; English.  
XX  
CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
CC useful for synthesizing sialylated oligosaccharide such as ganglioside,  
CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
CC other oligosaccharides that have biological activity. The enzymes and  
CC nucleic acids that encode them are useful for studies of the pathogenesis  
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
CC jejuni and the nucleic acids are used as probes to study expression of  
CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
CC against the glycosyltransferases are also useful for analyzing the  
CC expression patterns of these genes involved in pathogenesis. The nucleic  
CC acids are also useful for designing antisense oligonucleotides for  
CC inhibiting expression of the Campylobacter enzymes that are involved in  
CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
CC the host's immune system. The oligosaccharides are useful as diagnosing  
CC reagents or as therapeutics and as immunogens for producing antibodies.  
CC Bacterial glycosyltransferase can be used to catalyse the formation of  
CC oligosaccharides that are identical to the corresponding mammalian  
CC structures and are easier and less expensive to produce in large  
CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
CC origin of the enzymes facilitates expression of large quantities of the  
CC enzymes using relatively inexpensive prokaryotic expression systems.  
CC (Updated on 12-SEP-2003 to standardise OS field)  
XX  
SQ Sequence 291 AA;  
Query Match 99.6%; Score 1557; DB 3; Length 291;  
Best Local Similarity 99.7%; Pred. No. 9.6e-148;  
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLKCKKCAVFTPGFFFEQYY 60  
DB 1 MKKVIAGNGPSLKEIDYSRPNDFVRCNQFYEDKYLKCKKCAVFTPGFFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 120  
DB 61 TLKHLIQOEYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 120  
QY 121 PHEIYFNQIRITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 180  
DB 121 PHEIYFNQIRITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLPDFKN 180  
QY 181 DRSHYIGHSKNTDIALKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
DB 181 DRSHYIGHSKNTDIALKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHYPFGK 291  
DB 241 KDILIPSEAYGKFSKNINFKKIKIKENVYKLIKDLLRLPSDIKHYPFGK 291  
RESULT 3  
ID ABJ18482 standard; protein; 291 AA.  
XX  
AC ABJ18482;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni bifunctional sialtransferase cstII #4.  
XX  
KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAc transferase; N-Acetylgalactosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
KW ganglioside mimetics; inflammation; tumour metastasis.  
XX  
OS Campylobacter jejuni.

XX PN WO200274942-A2.  
XX PD 26-SEP-2002.  
XX PF 22-FEB-2002; 2002WO-CA000229.  
XX PR 21-MAR-2001; 2001US-00816028.  
XX PA (CANA ) NAT RES COUNCIL CANADA.  
XX PI Gilbert M, Wakarchuk WW;  
XX WPI: 2003-040554/03.  
XX N-PSDB; ABT13669.  
XX  
PT New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
PS Disclosure; Page 99; 107pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
CC Acetylglucosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present amino acid sequence represents a Campylobacter  
CC jejuni protein of the invention  
XX  
SQ Sequence 291 AA;  
Query Match 99.6%; Score 1557; DB 6; Length 291;  
Best Local Similarity 99.7%; Pred. No. 9.6e-148; Indels 0; Gaps 0;  
Matches 290; Conservative 0; Mismatches 1;  
QY 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYLGKCKKAVFTYTPGFFEQYY 60  
Db 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYLGKCKKAVFTYTPGFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291  
Db 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291  
RESULT 4  
ABJ18485  
ID ABJ18485 standard; protein; 291 AA.  
XX  
AC ABJ18485;  
XX  
DT 07-FEB-2003 (first entry)  
XX  
DE Campylobacter jejuni bifunctional sialtransferase cstII #6.

XX KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
KW GalNAC transferase; N-Acetylglucosamine transferase;  
KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
KW cytidine 5'-monophosphate sialic acid synthetase;  
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
XX ganglioside mimetics; inflammation; tumour metastasis.  
OS Campylobacter jejuni.  
XX  
XX WO200274942-A2.  
XX PD 26-SEP-2002.  
XX PF 22-FEB-2002; 2002WO-CA000229.  
XX PR 21-MAR-2001; 2001US-00816028.  
XX PA (CANA ) NAT RES COUNCIL CANADA.  
XX PI Gilbert M, Wakarchuk WW;  
XX WPI: 2003-040554/03.  
XX N-PSDB; ABT13671.  
XX  
PT New glycosyltransferases from Campylobacter, useful for synthesizing  
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
PT mechanisms of organisms that synthesize ganglioside mimetics.  
XX  
PS Disclosure; Page 99; 107pp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of  
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
CC Acetylglucosamine) transferase; galactosyltransferase;  
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
CC sequences of the invention are useful for ganglioside synthesis, studying  
CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
CC expression of Campylobacter enzymes involved in the biosynthesis of  
CC ganglioside mimetics that can mask the pathogen's from the host's immune  
CC system. The C. jejuni oligosaccharides of the invention may be used as  
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
CC metastasis). The present amino acid sequence represents a Campylobacter  
CC jejuni protein of the invention  
XX  
SQ Sequence 291 AA;  
Query Match 99.4%; Score 1553; DB 6; Length 291;  
Best Local Similarity 99.3%; Pred. No. 2.4e-147; Indels 0; Gaps 0;  
Matches 289; Conservative 0; Mismatches 2;  
QY 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYLGKCKKAVFTYTPGFFEQYY 60  
Db 1 MKKVIAGNGPSLKEIDYSRLPNDVDFRCNQFYFEDKYLGKCKKAVFTYTPGFFEQYY 60  
QY 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
Db 61 TLKHLIQOEYETELIMCSNYNQAHLENENFVKTFYDYPDAHLGYDFFKQLKEFNAYFK 120  
QY 121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
Db 121 FHEIYFNQRTSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDPKN 180  
QY 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
Db 181 DRSHYIGHSKNTDIKALEFLEKTYIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
QY 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291  
Db 241 KDILIPSEAYGKFSKNINFKIKIKENYVYKLIKDLLRLPSDIKHFFGK 291



XX SQ Sequence 291 AA;  
 Query Match 98.1%; Score 1534; DB 6; Length 291;  
 Best Local Similarity 98.3%; Pred. No. 2e-145;  
 Matches 286; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKKVIAGNGPSLKEIDYSLPNDPFRNCFYEDKYLKCKKAVFYTGFFPEQY 60  
 DB 1 MKKVIAGNGPSLKEIDYSLPNDPFRNCFYEDKYLKCKKAVFYTGFFPEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTVDYDPDAHLGYDFFKQKKEFNAYFK 120  
 DB 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTVDYDPDAHLGYDFFKQKKEFNAYFK 120  
 QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 DB 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKIKENIYVYKLIKDLLRLPSDIKHFKGK 291  
 DB 241 KDILIPSSAYGKFSKNINFKIKENIYVYKLIKDLLRLPSDIKHFKGK 291

RESULT 7  
 AAY97204  
 ID AAY97204 standard; protein; 291 AA.  
 XX AC AAY97204;  
 DT 12-SEP-2003 (revised)  
 DT 22-DEC-2000 (first entry)  
 XX DE Campylobacter jejuni OH4384 CstII sialyltransferase.  
 XX KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;  
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
 KW immunity; immunogen; ganglioside.  
 XX OS Campylobacter jejuni; OH4384.  
 XX PN WO200046379-A1.  
 XX PD 10-AUG-2000.  
 XX PF 01-FEB-2000; 2000WO-CA0000086.  
 XX PR 01-FEB-1999; 99US-0118213P.  
 XX PR 31-JAN-2000; 2000US-00495406.  
 XX PA (CANA ) NAT RES COUNCIL CANADA.  
 XX PI Gilbert M, Wakarchuk WW;  
 XX DR WPI; 2000-524418/47.  
 XX DR N-PSDB; AAA53721, AAA53720.  
 XX PT Novel glycosyltransferase polypeptides and polynucleotides useful for  
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 PT reagents and as immunogen for producing antibodies.  
 XX PS Claim 13; Page 91-92; 120pp; English.  
 XX CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
 CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and

CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC bacterial glycosyltransferase can be used to catalyse the formation of  
 CC oligosaccharides that are identical to the corresponding mammalian  
 CC structures and are easier and less expensive to produce in large  
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial  
 CC origin of the enzymes facilitates expression of large quantities of the  
 CC enzymes using relatively inexpensive prokaryotic expression systems.  
 CC (updated on 12-SEP-2003 to standardise OS field)  
 XX SQ Sequence 291 AA;  
 Query Match 97.5%; Score 1524; DB 3; Length 291;  
 Best Local Similarity 97.3%; Pred. No. 2e-144;  
 Matches 283; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MKKVIAGNGPSLKEIDYSLPNDPFRNCFYEDKYLKCKKAVFYTGFFPEQY 60  
 DB 1 MKKVIAGNGPSLKEIDYSLPNDPFRNCFYEDKYLKCKKAVFYTGFFPEQY 60  
 QY 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTVDYDPDAHLGYDFFKQKKEFNAYFK 120  
 DB 61 TLKHLIQOEYETELIMCSNYNOAHLENENFVKTVDYDPDAHLGYDFFKQKKEFNAYFK 120  
 QY 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 DB 121 FHEIYNQRIITSGVYMCVAIAIALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 DB 181 DRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYT 240  
 QY 241 KDILIPSSAYGKFSKNINFKIKENIYVYKLIKDLLRLPSDIKHFKGK 291  
 DB 241 KDILIPSSAYGKFSKNINFKIKENIYVYKLIKDLLRLPSDIKHFKGK 291  
 RESULT 8  
 ABJ18479  
 ID ABJ18479 standard; protein; 291 AA.  
 XX AC ABJ18479;  
 XX DT 07-FEB-2003 (first entry)  
 XX DE Campylobacter jejuni bifunctional sialtransferase cstII #1.  
 XX KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAc transferase; N-Acetylgalactosamine transferase;  
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX OS Campylobacter jejuni.  
 XX PN WO200274942-A2.  
 XX PD 26-SEP-2002.  
 XX PR 22-FEB-2002; 2002WO-CA0000229.  
 XX PR 21-MAR-2001; 2001US-00816028.



Db 181 DNSHYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEKNNYT 240  
 QY 241 KDILPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
 Db 241 KDILPSSSEAYGKFTKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 10  
 ABJ18481  
 ID ABJ18481 standard; protein; 291 AA.  
 XX  
 AC ABJ18481;  
 XX  
 DT 07-FEB-2003 (first entry)  
 XX  
 DE Campylobacter jejuni bifunctional sialtransferase cstII #3.  
 XX  
 KW Enzyme: Gene therapy; acyltransferase; glycosyltransferase;  
 KW GalNAC transferase; N-Acetylgalactosamine transferase;  
 KW Galactosyltransferase; sialyltransferase; sialic acid synthase;  
 KW cytidine 5'-monophosphate sialic acid synthetase;  
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;  
 KW ganglioside mimetics; inflammation; tumour metastasis.  
 XX  
 OS Campylobacter jejuni.  
 XX  
 PN WO200274942-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 22-FEB-2002; 2002WO-CA000229.  
 XX  
 PR 21-MAR-2001; 2001US-00816028.  
 XX  
 XX (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 PI Gilbert M, Wakarchuk WW;  
 XX  
 XX WPI; 2003-040554/03.  
 DR N-PSDB; ABT13668.  
 XX  
 XX New glycosyltransferases from Campylobacter, useful for synthesizing  
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis  
 PT mechanisms of organisms that synthesize ganglioside mimetics.  
 XX  
 PS Claim 5; Page 97-98; 107pp; English.  
 XX  
 XX The invention comprises the amino acid and coding sequences of  
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention  
 CC may be either an: acyltransferase; glycosyltransferase; GalNAC (N-  
 CC Acetylgalactosamine) transferase; galactosyltransferase;  
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)  
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein  
 CC sequences of the invention are useful for ganglioside synthesis, studying  
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit  
 CC expression of Campylobacter enzymes involved in the biosynthesis of  
 CC ganglioside mimetics that can mask the pathogen's from the host's immune  
 CC system. The C. jejuni oligosaccharides of the invention may be used as  
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour  
 CC metastasis). The present amino acid sequence represents a Campylobacter  
 CC jejuni protein of the invention  
 XX  
 SQ Sequence 291 AA;  
 Query Match 96.8%; Score 1513; DB 6; Length 291;  
 Best Local Similarity 95.9%; Pred. No. 2.6e-143;  
 Matches 279; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKKVIAGNPSLKEIDYSLPNDVFRNCQFYEDKYYLGKCKKAVFYTPGFPEQYY 60  
 Db 1 MKKVIAGNPSLKEIDYSLPNDVFRNCQFYEDKYYLGKCKKAVFYNPSPFPEQYY 60  
 QY 61 TLGHLIQNQEYETELIMCSNQNQAHLENENFVKTFDYFPDAHLGVDFFKQLKEFNAYFK 120

Db 61 TLGHLIQNQEYETELIMCSNQNQAHLENENFVKTFDYFPDAHLGVDFFKQLKEFNAYFK 120  
 QY 121 FHEIYFNQRTISGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 Db 121 FHEIYFNQRTISGVYMCVAIAALGYKEIYLSGIDFYQNGSSYAFDTKQENLLKLAPDFKN 180  
 QY 181 DRSHYIGHSKNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIQEKNNYT 240  
 Db 181 DNSHYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEKNNYT 240  
 QY 241 KDILPSSSEAYGKFSKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291  
 Db 241 KDILPSSSEAYGKFTKNINFKKIKENYVYKLIKDLLRLPSDIKHVFKGK 291

RESULT 11  
 AAY97215  
 ID AAY97215 standard; protein; 294 AA.  
 XX  
 AC AAY97215;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 22-DEC-2000 (first entry)  
 XX  
 DE Campylobacter jejuni NCTC11168 alpha-2,3-sialyltransferase.  
 XX  
 KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;  
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;  
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;  
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;  
 KW immunity; immunogen; ganglioside.  
 XX  
 OS Campylobacter jejuni; NCTC11168.  
 XX  
 PN WO200046379-A1.  
 XX  
 PD 10-AUG-2000.  
 XX  
 PF 01-FEB-2000; 2000WO-CA000086.  
 XX  
 PR 01-FEB-1999; 99US-0118213P.  
 PR 31-JAN-2000; 2000US-00495406.  
 XX  
 PA (CANA ) NAT RES COUNCIL CANADA.  
 XX  
 XX Gilbert M, Wakarchuk WW;  
 XX  
 XX WPI; 2000-524418/47.  
 XX  
 XX Novel glycosyltransferase polypeptides and polynucleotides useful for  
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic  
 PT reagents and as immunogen for producing antibodies.  
 XX  
 PS Disclosure; Page 97-98; 120pp; English.  
 XX  
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is  
 CC useful for synthesising sialylated oligosaccharide such as ganglioside,  
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for  
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and  
 CC other oligosaccharides that have biological activity. The enzymes and  
 CC nucleic acids that encode them are useful for studies of the pathogenesis  
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.  
 CC jejuni and the nucleic acids are used as probes to study expression of  
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised  
 CC against the glycosyltransferases are also useful for analyzing the  
 CC expression patterns of these genes involved in pathogenesis. The nucleic  
 CC acids are also useful for designing antisense oligonucleotides for  
 CC inhibiting expression of the Campylobacter enzymes that are involved in  
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from  
 CC the host's immune system. The oligosaccharides are useful as diagnosing  
 CC reagents or as therapeutics and as immunogens for producing antibodies.  
 CC Bacterial glycosyltransferase can be used to catalyse the formation of





PT New sialyltransferases useful for adding sialyl residues to acceptor  
 PT molecules.  
 XX  
 PS Claim 27; Fig 2; 47pp; English.  
 XX  
 CC The present sequence represents Campylobacter jejuni alpha-2,3-  
 CC sialyltransferase which is encoded by the cst-I gene. The alpha-2,3-  
 CC sialyltransferase protein is useful for producing desired carbohydrate  
 CC structures by contacting the acceptor molecule (which has a terminal  
 CC galactose residue) with an activated sialic acid molecule. The terminal  
 CC galactose residue is linked to a second residue (Glc or a GlcNAc, or  
 CC GlcNAc or GalNAc) in the acceptor molecule through a beta-1,3 or beta-1,4  
 CC linkage, respectively. The activated sialic acid is CMP-Neu5Ac. The  
 CC polynucleotides and polypeptides facilitate the improved production of  
 CC desired structures and nucleic acids encoding sialyltransferases  
 XX  
 SQ Sequence 430 AA;  
 Query Match 47.9%; Score 749; DB 2; Length 430;  
 Best Local Similarity 50.9%; Pred. No. 3.4e-66;  
 Matches 148; Conservative 45; Mismatches 80; Indels 18; Gaps 5;  
 QY 1 MKKVIAGNGPSLKEIDYSRLPNDVFRNQYFEDKYLGKCKKAVFYTPGFFPEQYV 60  
 Db 16 MQNIIAGNGSLKINIKRLPREYDVFRNQYFEDKYLGKCKKAVFYTPGFFPEQYV 75  
 QY 61 TLKHLIQNQYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAP 120  
 Db 76 TAKQILKNEVEIKNIFCSTFNLPFIESNDFLHQFYNFPPDAKLGVEIENLKEFYAYIK 135  
 QY 121 FHEIFYNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAP 177  
 Db 136 YNEIFYNKRITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAP 195  
 QY 178 FNDRSHYIGHKNTDIKALEFELEKTYKIKLYCLPCNSILLANFIELANLNSNFIIOEK- 236  
 Db 196 FAPNSC-----HSKEYDIEALKLLSIYKVNIIYALCDDSI LANHFFLSININNFTLENKH 251  
 QY 237 NNYTKDILIPSEAYGKFSKN-----INFKKI-KIKENVYKLIKDL 277  
 Db 252 NNSINDILLTNTPGVSYKQKADKNIMLFNFIHLHSKDNLIKFLNKEI 302  
 RESULT 14  
 ABR40195  
 ID ABR40195 standard; protein; 303 AA.  
 XX  
 AC ABR40195;  
 XX  
 DT 23-JUL-2003 (first entry)  
 XX  
 DE alpha-2,3/alpha-2,8-sialyltransferase #3.  
 XX  
 DE alpha-2,3/alpha-2,8-sialyltransferase; enzyme; sialic acid.  
 KW  
 XX Pasteurella multocida.  
 OS  
 XX Pasteurella multocida.  
 XX  
 PN WO2003027297-A1.  
 XX  
 PD 03-APR-2003.  
 XX  
 PF 26-SEP-2002; 2002WO-JP009907.  
 XX  
 XX 26-SEP-2001; 2001JP-00292796.  
 PR  
 XX (KYOM ) KYOMA HAKKO KOGYO KK.  
 PA  
 XX Endo T, Koizumi S;  
 PI  
 XX WPI; 2003-393339/37.  
 DR  
 XX N-PSDB; ACC71694.  
 XX  
 PT Alpha 2,3/alpha 2,8 sialyltransferase from Pasteurella multocida for

PT production of sialic acid-containing complex sugars.  
 XX  
 PS Claim 9; Page 60-62; 66pp; Japanese.  
 XX  
 CC The present invention relates to a method for producing alpha-2,3/alpha-  
 CC 2,8-sialyltransferase. alpha-2,3/alpha-2,8-sialyltransferase can be used  
 CC in a method for producing sialic acid-containing complex sugars from an  
 CC oligosaccharide substrate using a culture of a microorganism transformed  
 CC with DNA encoding alpha-2,3/alpha-2,8-sialyltransferase, or an extract of  
 CC the culture. The present sequence is a protein sequence for alpha-  
 CC 2,3/alpha-2,8-sialyltransferase from Pasteurella multocida  
 XX  
 SQ Sequence 303 AA;  
 Query Match 44.7%; Score 698.5; DB 6; Length 303;  
 Best Local Similarity 46.4%; Pred. No. 2.5e-61;  
 Matches 137; Conservative 58; Mismatches 93; Indels 7; Gaps 5;  
 QY 2 KKVIIAGNGPSLKEIDYSRLPNDVFRNQYFEDKYLGKCKKAVFYTPGFFPEQYV 61  
 Db 11 KAVIVAGNGESLSQIDYRLPKNYDVFRNQYFEDKYLGKCKKAVFYTPGFFPEQYV 70  
 QY 62 LKHLIQNQYETELIMCSNVAIAALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAP 120  
 Db 71 LVHLKRNNEYVDNVILSNFHTVDLEKSK-IQALFIDVINGYKLYSKLTAFDYLRL 129  
 QY 121 FHEIFYNQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAP 179  
 Db 130 YKELFENQRTITSGVVMCAVAIAALGYKEIYLSGIDFYQNGSSVAFDTKQENLLKLAP 189  
 QY 180 NDRSHYIGHKNTDIKALEFELEKTYKIKLYCLPCNSILLANFIELANLNSNFIIOEK 236  
 Db 190 KEKTLFYSYHSDIDLEALSFLQGHVHNFYSISPMSPSLSKHFPITVEDDCETTFAVPLK 249  
 QY 237 NNYTKDILIPSEAYGKFSKNINFKIKIKENVYKLIKDLPSDIKHVFKGK 291  
 Db 250 ENYINDILLPHFVVEKLGITVS-KKSRFHSNLIIVLIRDLKLPSALKHYLREK 303  
 RESULT 15  
 ADN19559  
 ID ADN19559 standard; protein; 958 AA.  
 XX  
 AC ADN19559;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Bacterial polypeptide #2212.  
 XX  
 KW Recombinant DNA construct; transformed plant; improved plant property;  
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
 KW pathogen tolerance; pest tolerance; plant disease resistance;  
 KW cell cycle pathway modification; plant growth regulator;  
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
 KW bacterial polypeptide.  
 XX  
 OS Bacteria.  
 XX  
 PN US2003233675-A1.  
 XX  
 PD 18-DEC-2003.  
 XX  
 PF 20-FEB-2003; 2003US-00369493.  
 XX  
 XX 21-FEB-2002; 2002US-0360039P.  
 XX  
 XX (CAOY/) CAO Y.  
 PA (HINK/) HINKLE G J.  
 PA (SLAT/) SLATER S C.  
 PA (CHEN/) CHEN X.  
 XX (GOLD/) GOLDMAN B S.  
 PA



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2006, 13:37:57 ; Search time 18.2031 Seconds  
(without alignments)  
1538.150 Million cell updates/sec

Title: US-10-734-719-12  
Perfect score: 1563  
Sequence: 1 MKKVIAGNGPSLKEIDYR.....KLIKOLLRLPSDIKHYFGK 291

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	804	51.4	294	2	G81318
2	485.5	31.1	231	2	E64149
3	122.5	7.8	706	2	D90124
4	117.5	7.5	590	2	G90127
5	117.5	7.5	621	2	E82912
6	117.5	7.5	960	2	S72284
7	115	7.4	707	2	T28418
8	114.5	7.3	2136	2	A05037
9	114	7.3	624	2	T28423
10	114	7.3	781	2	F90547
11	113.5	7.3	1272	2	H82926
12	112.5	7.2	440	2	F84955
13	112	7.2	347	2	T33018
14	111.5	7.1	1613	2	D90129
15	110.5	7.1	320	2	T28379
16	110	7.0	392	2	E65140
17	110	7.0	1084	2	C82931
18	109.5	7.0	908	2	F71613
19	109.5	7.0	1121	2	F71613
20	109.5	7.0	1532	2	T18438
21	109	7.0	839	2	T12827
22	109	7.0	1398	2	H71606
23	108.5	6.9	545	2	B70374
24	107.5	6.9	437	2	T28180
25	107	6.8	532	2	H30089
26	107	6.8	2269	2	T28677
27	106.5	6.8	590	2	E90089
28	106.5	6.8	1228	2	S37621
29	106.5	6.8	1284	2	G82897

alpha-N-acetylneur  
conserved hypothet  
exodeoxyribonuclea  
hypothetical prote  
hypothetical prote  
UDP-N-acetylmuramo  
hypothetical prote  
lipoprotein (impor  
N-myristoyltransfe  
hypothetical prote  
probable Zn-depend  
DNA/RNA helicase,  
hypothetical prote  
hypothetical prote  
probable secreted  
probable membrane

30 106 6.8 376 2 JC5600  
31 106 6.8 911 2 D90572  
32 106 6.8 971 2 A70179  
33 106 6.8 1021 2 E64576  
34 106 6.8 2539 2 E71619  
35 105.5 6.7 422 2 E64581  
36 105.5 6.7 504 2 T26424  
37 105.5 6.7 616 2 B90554  
38 104.5 6.7 341 2 G90112  
39 104.5 6.7 398 2 G81451  
40 104.5 6.7 439 2 A96981  
41 104.5 6.7 949 2 H97322  
42 104.5 6.7 1411 2 T18417  
43 104.5 6.7 1417 2 T18418  
44 104 6.7 1817 2 H71611  
45 104 6.7 2013 2 C71610

RESULT 1  
G81318  
hypothetical protein Cj1140 [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: G81318  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, A.P.; De Jong, P.; et al. 2003  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Van Vliet, A.; Whitehead, S.; Barrer, N.; et al. 2003  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals by  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: G81318  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-294 <PAR>  
A:Cross-references: UNIPROT:Q9PNF4; UNIPARC:UPI00000CA9C9; GB:AL139077; GB:AL111168; N  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj1140

Query Match 51.4%; Score 804; DB 2; Length 294;  
Best Local Similarity 53.9%; Pred. No. 4.7e-51;  
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

Qy 5 ITAGNGPSLKEIDYSLPNDPQVFRNCQFYFEDKYLYGKKCAVFTPGFFFEQYITLKH 64  
Db :::  
Qy 8 LVCGNGPSLKNIDYKSLPKQFQVFRNCQFYFEDKYLYGKKCAVFTPGFFFEQYITLKH 67  
Db :::  
Qy 65 LIQNGQYETELIMCSNNOAHLENFVKTFYDFPDHGLGDFKOLKEFNAYFKPHEI 124  
Db LIQNEYNENIVCSTINLEYIDGQFQVDFNFELYSDFDLGHEIILKLDKFAIKYNEI 127  
Qy 125 YFNQRTSGTYMCAVAIALGYKEIYLSGIDFYQNGSS-YAFDTKQENLLKLAPDFNDRS 183  
Db :::  
Qy 128 YNRQRTSGTYMCAVAIALGYKEIYLSGIDFYQNGSS-YAFDTKQENLLKLAPDFNDRS 187  
Db :::  
Qy 184 HYGHKNVDIKALEFLEKTYKIKYLCCLNSLLANFIELAPNLSNFIQEK-NNYTKD 242  
Db :::  
Qy 188 KFINHSMACDLQALDYLMKRYDNIYSL--NS--DEYFKLAPDYGSDVLSKKPKKYIND 243  
Db :::  
Qy 243 ILIPSESEA---YGFKEKSNINPKIKIKENNVYKLIKDLRLPSDIKHYFGK 291  
Db ILIPDKYAOERYYG-----KKSRLKENLHYKLDRLPSDIKHYLKEK 288

hypothetical protein HI0352 - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
C:Accession: E64149; S15287  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,

RESULT 2  
E64149  
hypothetical protein HI0352 - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 09-Jul-2004  
C:Accession: E64149; S15287  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage,

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630; PMID:7542800  
 A:Accession: E64149  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-231 <TIGR>  
 A:Cross-references: UNIPROT:P24324; UNIPARC:UPI0000139681; GB:U32720; GB:L42023; NID:g15  
 A:Experimental source: strain Rd KW20  
 R:Maekell, D.J.; Szabo, M.J.; Butler, P.D.; Williams, A.E.; Moxon, E.R.  
 Mol. Microbiol. 5, 1013-1022, 1991  
 A:Title: Molecular analysis of a complex locus from Haemophilus influenzae involved in p  
 A:Reference number: S15287; MUID:92065797; PMID:1956282  
 A:Accession: S15287  
 A:Molecule type: DNA  
 A:Residues: 1-231  
 A:Cross-references: UNIPARC:UPI000016782; EMBL:X57315; NID:g43587; PIDN:CAA40567.1; PID  
 A:Experimental source: strain RM7004  
 C:Superfamily: Haemophilus influenzae hypothetical protein HI0352

Query Match 31.1%; Score 485.5; DB 2; Length 231;  
 Best Local Similarity 43.7%; Pred. No. 3.9e-28;  
 Matches 101; Conservative 47; Mismatches 74; Indels 9; Gaps 6;  
 QY 65 LIQOEYE-TELIMCSNNOAHLENENFVKTFTYDYPDAHLGYDFPKOLKEFNAYKFEHE 123  
 Db 3 LIKNEYEYADILISSFVNLGDSLEKK-IRKVKQLLTVDIGHYVNLKLPAPDAYLQYNE 61  
 QY 124 IYFNQRTSGVYMCVAIAALGYKEIYLSGIDFYQ-NGSSYAFDTKQENLLKLAPDFKNDR 182  
 Db 62 LYENKRITSGVYMCVAVATVMGYKDLVLTGIDFYQEGNPNYAFHOKENIILKLPFSQNK 121  
 QY 183 SHYIGHSKNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLS--NFIIQEKNNYT 240  
 Db 122 SQSDIHSMEYDLNALLYFLQKHGYGNIYICISPEPLCNYPFLSP-LNNPITFLEEKQNT 180  
 QY 241 KDILIPSEAYGKFSKNINFKIKENYVYKIKDLRLPSDIKHVFKGK 291  
 Db 181 QDILIPKVFYKKTG---IYKPRYQNLIFRLWDILRLFNDKHALKS 228

RESULT 3  
 D90124  
 hypothetical protein rad3 [imported] - Guillardia theta nucleomorph  
 A:Species: nucleomorph Guillardia theta  
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: D90124  
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re  
 Nature 410, 1091-1096, 2001  
 A:Title: The highly reduced genome of an enslaved algal nucleus.  
 A:Reference number: A99082; MUID:11323671; PMID:11323671  
 A:Accession: D90124  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-706 <DOU>  
 A:Cross-references: UNIPROT:Q98S94; UNIPARC:UPI0000080CF5B; GB:AF083031; NID:g13794312; F  
 A:Genes: rad3  
 A:Map position: 3  
 A:Genome: nucleomorph  
 C:Keywords: nucleomorph

Query Match 7.8%; Score 122.5; DB 2; Length 706;  
 Best Local Similarity 22.0%; Pred. No. 0.26;  
 Matches 75; Conservative 56; Mismatches 111; Indels 99; Gaps 17;  
 QY 17 DYSRLPNDPFRNCNQYFEDKYLGKCAVFTPGFFEQYVTLKHLIQN-----69  
 Db 203 DYNFLILKFDQI-----IHSSKNYLNTRNKKIF----LLFSNFSFGALLKSKSIISLSL 254

QY 70 -----EYETELIMCSNNOAHLENENFVKTFTYDYPDAHLGYDF---KOLKEFNAYF-- 119  
 Db 255 KTIIDSYSRLIIVERDIKKKIFQINVLATFNFF-----HLKKNFFYIKKKNKILNNYLVLT 310  
 QY 120 -----KF-HEIYFNQRTSGVYMCVAIAALGYK--EIYLSGIDFYQNGSSYAFDTKQEN 170  
 Db 311 KLNFNQKFLHLYY---LNSYRFFLETLKLISKKEEYISSINIYKD-----IFLDEN 361  
 QY 171 LKLAP-----DFKNDRSHVIGHSKNTDIKALEFLEKTYKI-----KLYCLCPNSLLANFIE 222  
 Db 362 EIEIAPVVFQNFNDNLISHYKTYSPDNKSLNGIKKISNFIILTKFLLYQNNFINIKSE 421  
 QY 223 LAPN-----LNSNFIIOEKNNY--TKDILIPSSSE 249  
 Db 422 LDKNNSKFKLTLEIDFNFGIFEFFPELFRSIFITINDFEISKANNFFNCKNMFGLSGIK 481  
 QY 250 AYGKFSKNINFKIKI-----KENVY--YKLKDLRL 280  
 Db 482 NYQK-NKNCNYILTYVQNQNKASKKLDYEDYDFIGSILKL 521

RESULT 4  
 G90127  
 hypothetical protein orf590 [imported] - Guillardia theta nucleomorph  
 A:Species: nucleomorph Guillardia theta  
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: G90127  
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re  
 Nature 410, 1091-1096, 2001  
 A:Title: The highly reduced genome of an enslaved algal nucleus.  
 A:Reference number: A99082; MUID:11323671; PMID:11323671  
 A:Accession: G90127  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-590 <DOU>  
 A:Cross-references: UNIPROT:Q98S67; UNIPARC:UPI00000AB3FB; GB:AF083031; NID:g13794337; F  
 A:Genes: orf590  
 A:Map position: 3  
 A:Genome: nucleomorph  
 C:Keywords: nucleomorph

Query Match 7.5%; Score 117.5; DB 2; Length 590;  
 Best Local Similarity 21.9%; Pred. No. 0.49;  
 Matches 80; Conservative 53; Mismatches 108; Indels 125; Gaps 19;  
 QY 17 DYSRLPNDPFRNCNQF-----YFEDKYLGKCAVFTPGF 54  
 Db 5 DYLNLRLNLSVYSRCNFFTFVDLKKIFNFIQLNLINFFSEKLL-----51  
 QY 55 FFEQYVTLKHLIQ--NOEYETELIMCSNNOAHLE-NENFVK-----TFYDYF-----99  
 Db 52 IFONKFLHIYINRINEBLQFE-----KENEINFEKNDNYKSKFNLSLFDIFINEE 106  
 QY 100 -----PDHLGYDFPKOLKEFNAYKFEH-----EYFNQRTSGVY 135  
 Db 107 NFGILNFSKSLIFLDLNDKNIILSNFKIWLKNNFLKFNKLKILITTSFNSFSQSYFQ 166  
 QY 136 MCAVAIALGYK-----EYLSGI---DF---YONGSSYAFDTKQENLLKLAPDFKNDRSHVI 186  
 Db 167 FCNFAI--GSKIINFQVYKRIKNEFLSISDALSVFNKVDGLILILINEKFD-----219  
 QY 187 GHSKNNTDIKALEFLEKTYKLYCLCPNSLLANFIELAPNLSNFIIOEKNNYK-----241  
 Db 220 ---INELINSIYHISKIKKKKIFFP---PIINFI--MPDFVLNKIIRLKNVYKVICR 270  
 QY 242 -----DILIPSEAYGKFSKNINFKIKENYVYKIK-DLLRLPS-----DIK 285  
 Db 271 EYKSNFSLIFDLIVELFKSNEKYFKKINYNKTHENLNTKNNKFLKFLNLTFTTYTKWTILR 330  
 QY 286 HYFKGK 291

	Query Match	7.5%;	Score 117.5;	DB 2;	Length 960;
	Best Local Similarity	20.8%;	Pred. No.	0.87;	
	Matches	82;	Conservative	51;	Mismatches 107; Indels 155; Gaps 19;
Qy	27	VFCNCFYEDKYYLGG---KKCAVFTYPGFPEQQYTTLKHILQNQEYTELIIMCSNYNQ	83	:	:
Db	388	LYNNYNLPFYKYLQNQYIKCNFIYSISKNFK--YNLNII--IKYLLNVVIKYNSN	443	:	:
Qy	84	AHL-----ENENVKTIFYDPDAHLYGDFPKOLKEFNAYFK-----	120	:	:
Db	444	IQLLIKNIHNKWILVNIITY- - - - - YLYVHYHKFNLKYLNGKIILNMMNKYNIVFLI	496	:	:
Qy	121	- - - - - PHEIYFN-----		:	:
Db	497	NYNLFNSYNYKYKINNNNFINSNYVFCKMNFILKNFNIIQLINKLFYVNNIFIYYKYEK	556	:	:
Qy	148	--IYLSGD-----FYQGSSYAFD-----TKQENLLKLAPDFKNDRSHY--IGHSKNT	192	:	:
Db	557	KLFYLNINNIILKKYLYPNFYKTYTNKLPFIKYYNNFLYLYEIFKYNWYKYLLLNKYNKL	616	:	:
Qy	193	DIKALBFLEKTYX----IKLYCLCPNSSLANFIELAPLNSNFI-----IOEKN-	238	:	:

[illegible][illegible]

```
Db 422 ISYCNLNDNCLDNLNKLDELDSGNWEICFSSLPKSIKIKCKNCNITNTDFLNNLNL 481
Qy 159 -----GSSYAFDTKQENLLKLAPDFKNDRSHYIGHSKNTDIKALEFLEK--- 202
Db 482 QELDMSVNYININKSRIDNLPKLIKLA--NINNCDCAYISFKGLDNLELEMRNTCL 539
Qy 203 -----TYKI-----KLYCLC-----PNSL----- 216
Db 540 IDYLPKSLIKLHGTYHTFKIKDLNKLATLCLPGIINKINIEDLPKSLIELDISTCKYIINE 599
Qy 217 -----LANFIELAPNLNSNFIIO-----EKNNTKDI-----IPSSEAYG 252
Db 600 DCLKNLYNLKELTWSINCNIKQTLPSLIKIKLKCKYSYDDFKDIYDKKLNLNNEYG 659
Qy 253 KF-----SKNINFKKIKIKENYV-YKLIKDL 277
Db 660 KYISFLPEIIKKNYNAPKIHLDIYNLKKIKSI 693

RESULT 8
A05037
hypothetical protein 2136 - liverwort (Marchantia polymorpha) chloroplast
C:Species: chloroplast Marchantia polymorpha
C>Date: 05-Jun-1987, #sequence_revision 05-Jun-1987 #text_change 09-Jul-2004
C:Accession: S01591; A05037
R:Umesono, K.; Inokuchi, H.; Shiki, Y.; Takeuchi, M.; Chang, Z.; Fukuzawa, H.; Kohchi, T.
J. Mol. Biol. 203, 299-331, 1988
A:Title: Structure and organization of Marchantia polymorpha chloroplast genome. II. Gen
A:Reference number: S01567; MUID:89068686; PMID:2974085
A:Accession: S01591
A:Molecule type: DNA
A:Residues: 1-2136 <UME>
A:Cross-references: UNIPROT:P09975; UNIPARC:UPI000013A622; EMBL:X04465; NID:g11640; PIDN
R:Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shi
Nature 322, 572-574, 1986
A:Title: Chloroplast gene organization deduced from complete sequence of liverwort March
A:Reference number: A38014
A:Contents: annotation; gene organization, sites, features
C:Genetics:
A:Genome: chloroplast
C:Keywords: chloroplast

Query Match 7.3%; Score 114.5; DB 2; Length 2136;
Best Local Similarity 21.1%; Pred. No. 3.7;
Matches 74; Conservative 51; Mismatches 95; Indels 131; Gaps 15;

Qy 15 EIDYSLPNDPVDVRCNQFPEDKYILGKKCAVFTPGFFEQYITLKHILIQOEYETE 74
Db 829 ELFTSEINNDP-----LMRFFKLYIYRIYKDKELFNP-----IENRQLLQNFPEKTK 877
Qy 75 LIMCSNYNQAHLENENFVKTFDYDPDAHLGYDFFKQLKEFNAYFKFHEIYFNQRTSGV 134
Db 878 IL-----TFIDFLQDPENY-----NNRFHLEKTTIKNNLL 911
Qy 135 YMCVAIALGKYEIL-----SGIDFYONGSSYAFDT 166
Db 912 YLRLLKIFLKDKNFLLINEIKSFIEKKNNLFIKSQLSNVLLVKNYSKFFDNIENFHLK 971
Qy 167 KQENLLKLAPDFKNDRSHYIGHS--KNTDIKAL-----EFLEKTYIKLYCLCPNSL----- 216
Db 972 QKEKNIEIILNNQ-----YFEKSLLKTYLKNLNNNSKYSKYSKIFIFQLL--NILNKN 1026
Qy 217 -----LANFIELAPNLNSFIIO-----EKNNTKDIILIPSEAYGKFSKNINFKKIKIKE 267
Db 1027 YKTFQWISIELIFYSKNL--NYKIQNKIEKNYC-----YNNKISYKXKIKT 1071
Qy 268 -----NVY--YKLIKDLRLPSDTHKVFPGK 291
Db 1072 VNEFFKNNLFTNNNSWFFLEWWEYNTYIILQIIQETTFQITDVLVEYFKKK 1122
```

RESULT 9

```
T28423
ORF MSV261 leucine rich repeat gene family protein - Melanoplus sanguinipes entomopoxviri
C:Species: Melanoplus sanguinipes entomopoxvirus
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T28423
R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A:Reference number: Z20484; MUID:99102612; PMID:9847359
A:Accession: T28423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-624 <AFO>
A:Cross-references: UNIPROT:Q9YV11; UNIPARC:UPI00000EFC79; EMBL:AF063866; NID:g4049647;
C:Genetics:
A>Note: MSV261

Query Match 7.3%; Score 114; DB 2; Length 624;
Best Local Similarity 22.3%; Pred. No. 0.94;
Matches 69; Conservative 37; Mismatches 118; Indels 86; Gaps 15;

Qy 9 NGPSLKEIDYSR-----LPNDPVDVRCNQFPEDKYILGKKCAVFTPGFFEQY 60
Db 343 NLDNKLKELDISRNFNINISSLPKSLITLINSYIKHTN-----FINDLY 387
Qy 61 TLKHLIQOEYETEL-----IMCSNYNQAHLENENFVKTFDYDPDAHLGYDFFKQLKE 114
Db 388 NLKELDISRNFNINISSLPKSLIKLNSIDCKIKNTNFINDL-----DNLKELDI 436
Qy 115 FNAYPKFHEIYFNQRTSGVGVCAVAIALGKYEILSGIDFYONGSSYAFDTKQENLLKL 174
Db 437 SNYY-----NINISS-----LPKTLTTLNISSCNIKNTNFIN-----LDNLKE--LDI 478
Qy 175 APDFK-----NDRSHYIGHSKNT-----DIKALEFLEKTYIKLYCLCPNSLL 217
Db 479 SKNFNINISSLPKSLITLINSYINHTNFINDLYNLKELD--ISNNYNINISSL--PNTLI 536
Qy 218 ANFIELAPNLNSFIIOEKNNYTKDILIPSSSAYGKFSKNI---NPKKIKIKENV---Y 270
Db 537 ELNISSCYIKNTNFILNDN--LKKLEISKSLSDYMLHNIITFNVLKLLDKSNTSFTFY 594
Qy 271 YKLIKDLRL 280
Db 595 YKIKSKFVKI 604

RESULT 10
F90547
hypothetical protein MYPV 2860 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C:Accession: F90547
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: F90547
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-781 <KUR>
A:Cross-references: UNIPROT:Q98Q55; UNIPARC:UPI00000D45AF; GB:AL445566; PID:g14089700;
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV 2860
A:Genetic code: SGC3

Query Match 7.3%; Score 114; DB 2; Length 781;
Best Local Similarity 21.6%; Pred. No. 1.2;
Matches 71; Conservative 48; Mismatches 111; Indels 98; Gaps 16;

Qy 24 DFDVRCNQFY-----FEDKYILGKKCA-----VFYT 51
Db 134 DLSLVQONFYFTFRDKFNDTKLENPEFKYSAKALGSWEKLEKLTFSPEKFDLFLNS 193
```

QY 52 PGFFPEQYTLKHLIONQEYETELIMCSYNOAHLENENFVTFYDPPDAHLGYD----- 107  
 Db 194 YNFYESYNSQSLVINNEYR-----NYDK--LYNFYFEKQ-KDYRNKNVIDQDKIEL 243  
 QY 108 --FFKQLKEFNAYKPFHEIYFNQRTISGVYMCVAIALGYKEIY-----LSCIDFQNGS 160  
 Db 244 ATIRQIQLLYNKHKSSEIYKA-----KELYDNKKLLG-EFKK----- 282  
 QY 161 SYAFDTKQENLLKLAPDFKND--RSHYIGHSKNTD-----IKALEFFLEK-TYKIKL 208  
 Db 283 --SPATKKNVAVNMKSYINEFFLEYKLANSKAWDTTSNEEWSYKLTATVNMKISYKLLK 340  
 QY 209 YCLCNSLLANFIELAPNLNSNFIIOEKNNYTKDIL-----IPSSBAYGKFSKNINPKKI- 263  
 Db 341 VSFFKNLLNTNLLPFLKFNVEVRLIEQIKYKNEDEFIYFNLPKT-----FSKGFEKKLVL 395  
 QY 264 KIKENYVYKLIKDLRLPSDIKHYPK 291  
 Db 396 EIYDAYKPKIDKYLQRAQEHKNLYLKEK 423

RESULT 11  
 H82926  
 conserved hypothetical UU143 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: H82926  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 submitted to GenBank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
 A:Reference number: H82926  
 A:Status: preliminary  
 A:Accession: H82926  
 A:Molecule type: DNA  
 A:Residues: 1-1272 <GLA>  
 A:Cross-references: UNIPARC:UPI000000C1BB3; GB:AE002115; GB:AF222894; NID:G6899102; PIDN:  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: UU143  
 A:Genetic code: SGC3

Query Match 7.3%; Score 113.5; DB 2; Length 1272;  
 Best Local Similarity 24.9%; Pred. No. 2.4;  
 Matches 73; Conservative 33; Mismatches 88; Indels 99; Gaps 18;

QY 60 YTLKHLIONQEYETELI-----MCSYNOAHLE--NENFVKTFYD----- 98  
 Db 373 FKIKHINIDSSNKKAIIELDSSSSNDNTKLLLEANNKQLLIKSYYDNNPWSKIVSYEKKDN 432  
 QY 99 --FPDAHLGYDFFKQLKEFNAYKPFHEIYFNQRTISGVYMCVAIALGYKEIYLSGIDF 155  
 Db 433 NKMLFDLH--DFFPKDLKT-----FIITHIRFDDNITS-----LG--KIKENSPEY 473  
 QY 156 YONGSSYA-----FDTKQENLLKLAP--DFKNDRSHYIGHSKNTDIIKALEFFLEKTYKIK 207  
 Db 474 YQNDKEYLLKSLKYVFDIKENQLYGSACFNND-----DFKILK--NKTFFVK 520  
 QY 208 LYCLCPNSLLANFIELAPNLNSNFIIOEKN-----NY-----TKDILIP 246  
 Db 521 YEIDTKNMLNKYIPLNKYINVDV-----KNLAQFKIVNVFDGLNYKLESIKIVNKNLSLP 576  
 QY 247 SSEAYG-KFSKNINPK--KIKIKENYVK-----LIKDLRLPS-DIKHYFK 289  
 Db 577 YNDHVNIGNANNTFVSWHKYKPKQNIINEPFDSDSLIKDEFNLKIDFNLHMK 629

RESULT 12  
 F84955  
 UDP-N-acetylmuramoylalanine-D-glutamate ligase (EC 6.3.2.9) [imported] - Buchnera sp. (s  
 C:Species: Buchnera sp.  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 03-Jun-2002  
 C:Accession: F84955

R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
 Nature 407, 81-86, 2000  
 A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp.  
 A:Reference number: A84930; MUID:20445173; PMID:10993077  
 A:Accession: F84955  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-440 <STO>  
 A:Cross-references: UNIPARC:UPI000005E4D9; GB:AP000398; GSPDB:GN00144  
 A:Experimental source: strain APS  
 C:Genetics:  
 A:Gene: murD; BU218  
 C:Superfamily: UDP-N-acetylmuramate-alanine ligase  
 C:Keywords: ligase

Query Match 7.2%; Score 112.5; DB 2; Length 440;  
 Best Local Similarity 20.1%; Pred. No. 0.8;  
 Matches 70; Conservative 56; Mismatches 107; Indels 115; Gaps 19;

QY 3 KVIAGN--GPSLKEIDYSRLPNDVFRNCNQFYFEDKYILGKKCAVF-----YT 51  
 Db 134 KAFLLGNGVGVLEILDKEA---DLVIELSSFQLENTFNLSKIAVILNISSEHINRYP 190  
 QY 52 PGFFPEQYTLKHLIONQEYETELIMCSYNO-----AHLENENFVTFYDPPDAHLGY 106  
 Db 191 NG--FOQYKNTKLSVYNQ--ABICINSNDKIEKSLIHSKNKKWIS-----FGTNRSDY 240  
 QY 107 DFFKQLKEFNAYKPFHEIYFNQRTISGVYMCVAIALGYKEIYLSGIDFYON-----GSS 161  
 Db 241 RICSNSNDPILFFK-----NKKI-----LNTSEILLYGVHNNILVSLAIS 282  
 QY 162 YAFDTKQENLLKLAPDF-----KNDRS-HYIGHSKNTDIIKA----- 196  
 Db 283 DAMQPRNDAINVLKSFSLPHRFQIIKNEKGVYRINDSKSTNVNSTQVALNSIKTTGTI 342  
 QY 197 -----LEFLEKTYKIKLYC-----LCPSLLANFIELAPNLNSNF 231  
 Db 343 RLLGGDSKSNANFNILKNTFKIKIYCFGRDGIKLSKICEKSI--YVE--NLKKAV 397  
 QY 232 IIOEKNNYTKD--ILIPSEAYGKFSKNINFKIKIKIKENYVYKLIKDL 277  
 Db 398 ILISKQVSGDVTLLSPGCSSILGQFS---NFEE---RGNLFILKIKEI 439

RESULT 13  
 T33018  
 hypothetical protein K07H8.8 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T33018  
 R:Pulton, B.; Hawkins, J.; Gattung, S.; Wohldmann, P.; Elliott, G.  
 submitted to the EMBL Data Library, February 1998  
 A:Description: The sequence of C. elegans cosmid K07H8.  
 A:Reference number: Z21264  
 C:Genetics:  
 A:Accession: T33018  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-347 <FUL>  
 A:Cross-references: UNIPROT:Q45177; UNIPARC:UPI0000079457; EMBL:AF047659; PIDN:AAC04429  
 A:Experimental source: strain Bristol N2; clone K07H8  
 C:Genetics:  
 A:Gene: CESP.K07H8.8  
 A:Map position: 4  
 A:Introns: 47/1; 110/3; 158/2; 202/2; 249/2; 302/1

Query Match 7.2%; Score 112; DB 2; Length 347;  
 Best Local Similarity 20.5%; Pred. No. 0.66;  
 Matches 71; Conservative 54; Mismatches 108; Indels 114; Gaps 17;

QY 7 AGNGPSLKEID-YSLPNDVFRNCNQFYFEDKYILGKK-----CKAVYTFGPF-----F 56  
 Db 15 SGEAGDIDDLHFTMLNPFY---CTT---SNKFYVYKERYIPKAVLVVWVGIFSI CFSI 67

Qy	57	BOYTLKHLIQEYETELIMCSNVNQAHLENENFKTYDYFPD-----	101
Db	68	DKYQLMKESKCNVDLAECNNLP-----FNQTFVPPFYDLOFTTISPRYDISLCLIP	123
Qy	102	---AHLGYDFFQKLKEFNAYKPFHEIYFNQRITSGVWYCAVAIALGYKEIYLSGIDFYON	158
Db	124	KVSTIGTAICYIQDPEAFYK-----NNRTIS-----TEMY---GGFCFN	162
Qy	159	GSSVAPDTKQENLKLAPDFKNDRSHY-----IGHSKNTDIKALEPL	200
Db	163	NELKSFWMQVRSLSNSFDNIIFTRNPFYDRFISGFTFEKCVNNTDTNICHGCGTDIRC--PL	220
Qy	201	EKTYK--IKLYCLCPNSLLAN-----FTELAPLNSNFTIOEKNNYTKDIL-----	244
Db	221	QKEYRLLIRMSMLFPVYTGADTHFAPQWYCDMKNNIKNSTVQVSSCTEKKIMDNLL	280
Qy	245	-----IPSE-----AYGESKNINFKKIKENVYVKLND	276
Db	281	TVFKNRNVPEENLNEVREELSGKGTQHSSTGSLRLK---YEKLTEE	324

RESULT 14

D90129 hypothetical protein orf1613 [imported] - Guillardia theta nucleomorph  
D90129 C;Species: nucleomorph Guillardia theta  
C;Note: a nucleomorph of the vestigial nucleus of a eukaryotic endosymbiont  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C;Accession: D90129  
R;Doughlas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reid,  
Nature 410, 1091-1096, 2001  
A;Title: The highly reduced genome of an enslaved algal nucleus.  
A;Reference number: A99082; PMID:11323671; PMID:11323671  
A;Accession: D90129  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1613 <DOU>  
A;Cross-references: UNIPROT: Q98S55; UNIPARC:UPI000000A244C; GB:AF083031; NID:gl3794350; F:  
C;Genetics:  
A;Gene: orf1613  
A;Map position: 3  
A;Genome: nucleomorph  
C;Keywords: nucleomorph

Query Match	7.1%;	Score 111.5;	DB 2;	Length 1613;
Best Local Similarity	21.8%;	Pred. No. 4.4;		
Matches	77;	Conservative 45;	Mismatches 101;	Indels 131; Gaps 19;
Qy	16	IDYSRLPNDFVRCNQFYEDKYYLKKC---	KAVETPTGFFPEQYVYTLKHLIQEYE	72
Db	1302	IKUNKIYDRDKTILHTIYVSVNYFIKKCDKVMLEFYK	-----FLMKRIFLDS--	1350
Qy	73	TELIMCSNYQNAHLENENF-----VKTFDYDFPDAHL	-----GYDFPKQL	112
Db	1351	-----VNYNY-KISYEIFFEQISVVISFNSLPKTNLTKMKNLP	FIKKFNLIINLIILKEL	1403
Qy	113	KFNNAVY-----KTH-----EIV-FNQRITSGVMYCAVAIALGYKEIYLSGIDFYQN	158	
Db	1404	RNFSATFLIILLKHEFLNINSIELLKFNFYQKFNFIYLI	-----FSKN	1446
Qy	159	GSSVAFDTQENILLKAPDPKNDORSYVIGHSKNTDIK	-----ALEFLE-----	201
Db	1447	GSSFI-----SLHUKKFLUPENKYQSKILS-SKYVTKIFFPEILYFFIQSPFKLELNFPS	1501	
Qy	202	-----KTY--KIKLYCLCP-NSLIANFIELAPNLSNFI-IQEKQNYTKD	242	
Db	1502	NFNWLNKKELLYFTKQYLLKLIYSIPTNLLIKQINVLKULHLEFISIQ	-----	1554
Qy	243	ILIPSSEAYGKFSKNINFKKIKIKENYVYKLIK	-----DLLRUPSDIKHY	287
Db	1555	--ILKKKLFNMLKNKFIFYRAKDYNSYFSFKTKRKFIFYNSLGFINTKTDIENF	1606	

RESULT 15

T28379  
ORF MSV218 hypothetical protein - Melanoplus sanguinipes entomopoxvirus  
C:Species: Melanoplus sanguinipes entomopoxvirus  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T28379  
R:Afonso, C.L.; Tushman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.  
J. Virol. 73, 533-552, 1999  
A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.  
A:Reference number: Z20484; MUID:99102612; PMID:9847359  
A:Accession: T28379  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-320 <AFO>  
A:Cross-references: UNIPROT:Q9YVM4; UNIPARC:UPI000000EF5DD; EMBL:AF063866; NID:g  
C:Genetics:  
A:Note: MSV218

Query Match	7.1%	Score 110.5;	DB 2;	Length 320;
Best Local Similarity	25.0%;	Pred. No. 0.77;		
Matches	77;	Conservative 37;	Mismatches 85;	Indels 109; Gaps 22;
Qy	40	YLCKK---CKAVFYTPGFPEQYYTLKH-----LQIQEYETEYELIMCSNYNQAHLNEN	90	
Db	2	YLKSNICTSQVINNP-----LYNIKOSKKVCIHKNYYK--ILCENY---HLFNKS	50	
Qy	91	-FVKTFDY-----FPDAHLGYDFEKKQKFENAYFPHEIFYFNORITSGVYMCVAVALG	144	
Db	51	RNIKKVFDYILKNIFPLSLESIIYINNELSKVINHIEVLYREY-----	95	
Qy	145	YKBIIYLGIDFYQNG-----SSYAFDTKQENL-----LKLAPDF---KNDRS	183	
Db	96	YTKFLNSNELDENHNLWINLYLSTSEYNFNTDFNNIKQLQFNPLKIL-DLNI	PKND--	152
Qy	184	HYTGHKNVTDIKALEPLEK-TYIKLYLCYICPSLLANFELAPLNSNFIQEKNYT--	240	
Db	153	---HNKS---GILFKNITNYKK-----PNN-----YLLYSOGIIIFYNPEKNCCSN	194	
Qy	241	-----KQILIP-SSEAYGKFSKNINFKKI-----KIKENVYVYIKLRLRP	281	
Db	195	LNNIIKFKEISIELASILG---KKINYKNIQWALYIMYKSKKILKNYKLLKILFPNT	251	
Qy	282	SDIKHYFK	289	
Db	252	YD-EYVFK	258	

Search completed: July 19, 2006, 13:48:11  
Job time : 20.2031 secs



GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 16:14:03 ; Search time 207 Seconds  
(without alignments)

3945.599 Million cell updates/sec

Title: US-10-734-719-9

Perfect score: 1563

Sequence: 1 MKKVIIAGNPSLKEIDYSR.....KLIKDLRLPSDIKHVFKGK 291

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp  
-Q=/abs/ABSSWEB.spool/US10734719/runat\_21082006\_142143\_12958/app\_query.fasta\_1  
-DB=Issued Patents NA -QFIX=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=1000 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=150  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs803p -USER=US10734719 @CCN 1.1 307 @runat\_21082006\_142143\_12958  
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /EMC Celerra\_SIDS3/prodata/2/ina/1 COMB.seq.\*  
2: /EMC Celerra\_SIDS3/prodata/2/ina/5 COMB.seq.\*  
3: /EMC Celerra\_SIDS3/prodata/2/ina/6A COMB.seq.\*  
4: /EMC Celerra\_SIDS3/prodata/2/ina/6B COMB.seq.\*  
5: /EMC Celerra\_SIDS3/prodata/2/ina/7 COMB.seq.\*  
6: /EMC Celerra\_SIDS3/prodata/2/ina/H COMB.seq.\*  
7: /EMC Celerra\_SIDS3/prodata/2/ina/FACTUS COMB.seq.\*  
8: /EMC Celerra\_SIDS3/prodata/2/ina/PP COMB.seq.\*  
9: /EMC Celerra\_SIDS3/prodata/2/ina/RE COMB.seq.\*  
10: /EMC Celerra\_SIDS3/prodata/2/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	876	3	US-09-495-406-8
2	1563	100.0	876	3	US-09-816-028A-8
3	1563	100.0	876	3	US-10-303-162-8
4	1563	100.0	876	3	US-10-303-134-8
5	1563	100.0	876	3	US-10-303-118-8
6	1563	100.0	876	3	US-10-303-128-8
7	1563	100.0	876	5	US-10-735-419-8
8	1559	99.7	873	3	US-09-816-028A-13

9	1559	99.7	873	3	US-10-303-162-13	Sequence 13, Appl
10	1559	99.7	873	3	US-10-303-134-13	Sequence 13, Appl
11	1559	99.7	873	3	US-10-303-118-13	Sequence 13, Appl
12	1559	99.7	873	3	US-10-303-128-13	Sequence 13, Appl
13	1559	99.7	873	5	US-10-735-419-13	Sequence 13, Appl
14	1557	99.6	873	3	US-09-816-028A-11	Sequence 11, Appl
15	1557	99.6	873	3	US-10-303-162-11	Sequence 11, Appl
16	1557	99.6	873	3	US-10-303-134-11	Sequence 11, Appl
17	1557	99.6	873	3	US-10-303-118-11	Sequence 11, Appl
18	1557	99.6	873	3	US-10-303-128-11	Sequence 11, Appl
19	1557	99.6	873	5	US-10-735-419-11	Sequence 11, Appl
20	1528	97.8	876	3	US-09-495-406-4	Sequence 4, Appl
21	1528	97.8	876	3	US-09-816-028A-4	Sequence 4, Appl
22	1528	97.8	876	3	US-10-303-162-4	Sequence 4, Appl
23	1528	97.8	876	3	US-10-303-134-4	Sequence 4, Appl
24	1528	97.8	876	3	US-10-303-118-4	Sequence 4, Appl
25	1528	97.8	876	3	US-10-303-128-4	Sequence 4, Appl
26	1528	97.8	876	5	US-10-735-419-4	Sequence 4, Appl
27	1525	97.6	876	3	US-09-495-406-2	Sequence 2, Appl
28	1525	97.6	876	3	US-09-816-028A-2	Sequence 2, Appl
29	1525	97.6	876	3	US-10-303-162-2	Sequence 2, Appl
30	1525	97.6	876	3	US-10-303-134-2	Sequence 2, Appl
31	1525	97.6	876	3	US-10-303-118-2	Sequence 2, Appl
32	1525	97.6	876	3	US-10-303-128-2	Sequence 2, Appl
33	1525	97.6	876	5	US-10-735-419-2	Sequence 2, Appl
34	1525	97.6	876	3	US-09-495-406-1	Sequence 1, Appl
35	1525	97.6	876	3	US-09-816-028A-1	Sequence 1, Appl
36	1525	97.6	876	3	US-10-303-162-1	Sequence 1, Appl
37	1525	97.6	876	3	US-10-303-134-1	Sequence 1, Appl
38	1525	97.6	876	3	US-10-303-118-1	Sequence 1, Appl
39	1525	97.6	876	3	US-10-303-128-1	Sequence 1, Appl
40	1525	97.6	876	5	US-10-735-419-1	Sequence 1, Appl
41	1514	96.9	876	3	US-09-495-406-6	Sequence 6, Appl
42	1514	96.9	876	3	US-09-816-028A-6	Sequence 6, Appl
43	1514	96.9	876	3	US-10-303-162-6	Sequence 6, Appl
44	1514	96.9	876	3	US-10-303-134-6	Sequence 6, Appl
45	1514	96.9	876	3	US-10-303-118-6	Sequence 6, Appl
46	1514	96.9	876	3	US-10-303-128-6	Sequence 6, Appl
47	1514	96.9	876	5	US-10-735-419-6	Sequence 6, Appl
48	743	47.5	1293	3	US-09-272-960-1	Sequence 1, Appl
49	743	47.5	1293	3	US-10-058-636-1	Sequence 1, Appl
50	721.5	46.2	1830121	3	US-09-557-884-1	Sequence 1, Appl
51	721.5	46.2	1830121	3	US-09-643-990A-1	Sequence 1, Appl
52	721.5	46.2	1830121	3	US-10-158-865-1	Sequence 1, Appl
53	113.5	7.3	2375	3	US-09-113-750A-33	Sequence 33, Appl
54	113	7.2	32392	3	US-09-662-254B-27	Sequence 27, Appl
55	113	7.2	640681	3	US-09-790-988-1	Sequence 1, Appl
56	112.5	7.2	580073	3	US-08-545-528D-1	Sequence 1, Appl
57	112	7.2	1179	3	US-09-492-709A-188	Sequence 188, App
58	112	7.2	4810	3	US-08-852-629-11	Sequence 11, Appl
59	112	7.2	4838	3	US-08-852-629-15	Sequence 15, Appl
60	112	7.2	50000	3	US-09-662-254B-25	Sequence 25, Appl
61	111.5	7.1	1866	3	US-09-601-198-153	Sequence 153, App
62	110.5	7.1	11474	3	US-09-495-406-1	Sequence 1, Appl
63	110.5	7.1	11474	3	US-09-816-028A-1	Sequence 1, Appl
64	110.5	7.1	11474	3	US-10-303-162-1	Sequence 1, Appl
65	110.5	7.1	11474	3	US-10-303-134-1	Sequence 1, Appl
66	110.5	7.1	11474	3	US-10-303-118-1	Sequence 1, Appl
67	110.5	7.1	11474	3	US-10-303-128-1	Sequence 1, Appl
68	110.5	7.1	11474	5	US-10-735-419-1	Sequence 1, Appl
69	110.5	7.1	50000	3	US-09-662-254B-23	Sequence 23, Appl
70	110	7.0	580073	3	US-08-545-528D-1	Sequence 1, Appl
71	108	6.9	2469	3	US-09-662-254B-32	Sequence 32, Appl
72	108	6.9	50000	3	US-09-662-254B-24	Sequence 24, Appl
73	107.5	6.9	888	3	US-09-107-532A-2598	Sequence 2598, Ap
74	107.5	6.9	1182	3	US-09-601-198-18	Sequence 18, Appl
75	107.5	6.9	50000	3	US-09-662-254B-23	Sequence 23, Appl
76	106	6.8	4152	3	US-09-662-254B-9	Sequence 9, Appl
77	106	6.8	50000	3	US-09-662-254B-25	Sequence 25, Appl
78	105.5	6.7	6768	2	US-08-107-755A-1	Sequence 1, Appl
79	105.5	6.7	8457	2	US-07-991-867B-1	Sequence 1, Appl
80	105.5	6.7	8457	2	US-08-544-332-1	Sequence 1, Appl
81	105.5	6.7	8457	3	US-09-370-861A-1	Sequence 1, Appl















958	77	4.9	1650	3	US-09-166-460-10	Sequence 10, Appl
959	77	4.9	1650	3	US-09-166-460-12	Sequence 12, Appl
960	77	4.9	1650	3	US-09-166-460-14	Sequence 14, Appl
961	77	4.9	1650	3	US-09-166-460-18	Sequence 18, Appl
962	77	4.9	1650	3	US-09-361-718-4	Sequence 4, Appl
963	77	4.9	1650	3	US-09-361-718-6	Sequence 6, Appl
964	77	4.9	1650	3	US-09-361-718-10	Sequence 10, Appl
965	77	4.9	1650	3	US-09-361-718-12	Sequence 12, Appl
966	77	4.9	1650	3	US-09-361-718-14	Sequence 14, Appl
967	77	4.9	1650	3	US-09-361-718-18	Sequence 18, Appl
968	77	4.9	1650	3	US-09-860-282B-4	Sequence 4, Appl
969	77	4.9	1650	3	US-09-860-282B-6	Sequence 6, Appl
970	77	4.9	1650	3	US-09-860-282B-10	Sequence 10, Appl
971	77	4.9	1650	3	US-09-860-282B-12	Sequence 12, Appl
972	77	4.9	1650	3	US-09-860-282B-14	Sequence 14, Appl
973	77	4.9	1650	3	US-09-860-282B-18	Sequence 18, Appl
974	77	4.9	1650	5	US-09-974-300-1675	Sequence 1675, Ap
975	77	4.9	1785	3	US-09-830-230A-376	Sequence 376, App
976	77	4.9	1827	3	US-08-956-171B-450	Sequence 450, App
977	77	4.9	1827	3	US-08-781-986A-450	Sequence 450, App
978	77	4.9	1878	3	US-09-830-230A-375	Sequence 375, App
979	77	4.9	1959	3	US-09-166-460-1	Sequence 1, Appl
980	77	4.9	1959	3	US-09-361-718-1	Sequence 1, Appl
981	77	4.9	1959	3	US-09-398-395A-27	Sequence 27, Appl
982	77	4.9	1959	3	US-09-887-586A-27	Sequence 27, Appl
983	77	4.9	1959	3	US-09-895-752-27	Sequence 27, Appl
984	77	4.9	1959	3	US-09-903-012B-27	Sequence 27, Appl
985	77	4.9	1959	3	US-09-900-797-27	Sequence 27, Appl
986	77	4.9	1959	3	US-09-860-282B-1	Sequence 1, Appl
987	77	4.9	1959	3	US-09-893-820-27	Sequence 27, Appl
988	77	4.9	1968	3	US-09-601-198-100	Sequence 100, App
989	77	4.9	2193	3	US-09-398-865A-1	Sequence 1, Appl
990	77	4.9	2193	3	US-09-710-714-1	Sequence 1, Appl
991	77	4.9	2628	3	US-09-255-829-9	Sequence 9, Appl
992	77	4.9	3108	3	US-09-614-221A-194	Sequence 194, App
993	77	4.9	3183	3	US-09-248-796A-2521	Sequence 2521, Ap
994	77	4.9	3186	2	US-08-701-846-1	Sequence 1, Appl
995	77	4.9	3282	3	US-09-315-793-51	Sequence 51, Appl
996	77	4.9	3763	2	US-07-792-865D-1	Sequence 1, Appl
997	77	4.9	3795	3	US-09-710-279-3632	Sequence 3632, Ap
998	77	4.9	3950	3	US-09-398-395A-33	Sequence 33, Appl
999	77	4.9	3950	3	US-09-887-586A-33	Sequence 33, Appl
1000	77	4.9	3950	3	US-09-895-752-33	Sequence 33, Appl

US-09-495-406-8

Alignment Scores:

Pred. No.: 3,66e-173

Score: 1563.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Query Match: 100.0%

DB: 3

US-10-734-719-9 (1-291) x US-09-495-406-8 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20

Db 1 ATGAAAAAGTTTATTATGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

Db 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

Db 121 CTTGGTAAAAATGCAAAAGCAGTCTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80

Db 181 ACTTTAAACATTTTAAATCCAAAAATCAAGAAATATGAGCCGAACTAATTTATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100

Db 241 TACACCAAGCTCATCTAGAAAAATGMAAATTTGTAAAAACTTTTACGATTTATTTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

Db 301 GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTGACGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATACCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180

Db 481 TCTTATGCTTTTGATACCAAAACAGAAATCTTTTAAAACTAGCCCTCGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

Db 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220

Db 601 GAAAAAATTACAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240

Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATACTACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

Db 721 AAAGATATACTCATACCTTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280

Db 781 AAAAAATATAAATTAAGAAATGTTTATTAACAGTTTCATATAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291

Db 841 CCTAGTATATAAGCAATTTATTTCAAAGGAAAA 873

RESULT 2

US-09-495-406-8

Sequence 8, Application US/09495406

Patent No. 6503744

GENERAL INFORMATION:

APPLICANT: Gilbert, Michel

APPLICANT: Makarchuk, Warren W.

APPLICANT: National Research Council of Canada

TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

TITLE OF INVENTION: Gangliosides and Ganglioside Mimics

FILE REFERENCE: 019633-000110US

CURRENT APPLICATION NUMBER: US/09/495,406

CURRENT FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: US 60/118,213

PRIOR FILING DATE: 1999-02-01

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 8

LENGTH: 876

TYPE: DNA

ORGANISM: Campylobacter jejuni

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(876)

OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II

OTHER INFORMATION: (cstII) from C. jejuni O:19





```
Db 61 CTACCAAAATGATTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGTGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuLeuGlnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAATAATATGTTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluGlnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgLleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGCAAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyLleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTCATTTTATCAAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrLleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysLleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAAAATAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheLleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATACTACTACT 720
Qy 241 LysAspLleLeuLleProSerSerGluAlaTyrGlyLysPheSerLysAsnLleAsnPhe 260
Db 721 AAAGATATACTATACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysLleLysLysGluAsnValTyrTyrLysLeuLleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAAATTAAGAAAAATGTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCAATTTATTTCAAGGAAAA 873
```

## RESULT 4

```
US-10-303-134-8
; Sequence 8, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; ; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
```

```
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni 0.19
US-10-303-134-8
```

```
Alignment Scores:
Pred. No.: 3,66e-173 Length: 876
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-9 (1-291) x US-10-303-134-8 (1-876)

```
Qy 1 MetLysLysValLleLleAlaGlyAsnGlyProSerLeuLysGluLleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATCATTTTGTATGATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTTACACCCCTAAATTTTCTTTTGTGCAATACTACT 180
Qy 61 ThrLeuLysHisLeuLleGlnAsnGlnGluTyrGluThrGluLeuLleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATTCAAAAATCAAGAAATATGAGACCGAATAATATGTTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATATTTTCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgLleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTGCAAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLleLysLeuSerGlyLleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrLleGlyHisSerLysAsnThrAspLleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAATAACAGATATATAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysLleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAAAATAAACTATATTTGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheLleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATACTACTACT 720
Qy 241 LysAspLleLeuLleProSerSerGluAlaTyrGlyLysPheSerLysAsnLleAsnPhe 260
```

```
Db 721 AAAGATATCTCATACCTTCTAGTGAGGCTTATCGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLysLysLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTTGATAAAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 5
US-10-303-118-8
; Sequence 8, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-118-8

Alignment Scores:
Pred. No.: 3,66e-173 Length: 876
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-118-8 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTATGCTGGAAATGGACCAAGTTTAAAGAAATATGATTATTCAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTGAGATAATACTAT 120
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAAATGCAAGACAGTGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAAATCAAGAATATGAGACCGAACTAATATATGTTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGAAAAAATTTTACGATTTATTTTCCT 300
```

```
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGTGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAACTTACAAAAATAAACTATATTCCTTATGCTCTAAATAGTCTTTTAGCAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTTATCATACAGAAAAATAACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAAAGTTTGATAAAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 6
US-10-303-128-8
; Sequence 8, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:19
US-10-303-128-8

Alignment Scores:
```

Pred. No.:		3.66e-173	Length:	876
Score:		1563.00	Matches:	291
Percent Similarity:		100.0%	Conservative:	0
Best Local Similarity:		100.0%	Mismatches:	0
Query Match:		100.0%	Indels:	0
DB:		3	Gaps:	0
US-10-734-719-9 (1-291) x US-10-303-128-8 (1-876)				
QY	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20	
DB	1	ATGAAAAAGTTATTATGCTGGAATGGCAAGTTTAAAGAAATTTGATTATTCAGG	60	
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40	
DB	61	CTACCAATGATTTGATGTTAGTGTATCAATTTTATTTTGAAGATAAATACTAT	120	
QY	41	LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60	
DB	121	CTTGTAAAAATGCAAGCAAGTGTATTTACACCCCTAATTTCTTTGAGCAATAC	180	
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80	
DB	181	ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATTTATGTGTCTAAT	240	
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100	
DB	241	TACACCAAGCTCATCTAGAAAAATTTTGAAGAACTTTTACGATTATTTTCCT	300	
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120	
DB	301	GATGCTCATTTGGGATATGATTTTTTAAACACTTAAAGAAATTTAATGCTTATTTTAA	360	
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140	
DB	361	TTTCAGGAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC	420	
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160	
DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA	480	
QY	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180	
DB	481	TCCTATGCTTTTGTATACCAAAAGAAATCTTTTTTAAAAAATCTAGCCCTGATTTTAA	540	
US-10-734-719-9 (1-291) x US-10-735-419-8 (1-876)				
QY	1	MethLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20	
DB	1	ATGAAAAAGTTATTATGCTGGAATGGCAAGTTTAAAGAAATTTGATTATTCAGG	60	
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40	
DB	61	CTACCAATGATTTGATGTTAGTGTATCAATTTTATTTTGAAGATAAATACTAT	120	
QY	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60	
DB	121	CTTGTAAAAATGCAAGCAAGTGTATTTACACCCCTAATTTCTTTGAGCAATAC	180	
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80	
DB	181	ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAACTAATTTATGTGTCTAAT	240	
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100	
DB	241	TACACCAAGCTCATCTAGAAAAATTTTGAAGAACTTTTACGATTATTTTCCT	300	
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120	
DB	301	GATGCTCATTTGGGATATGATTTTTTAAACACTTAAAGAAATTTAATGCTTATTTTAA	360	
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140	
DB	361	TTTCAGGAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTAGCC	420	
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160	
DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA	480	
QY	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180	
DB	481	TCCTATGCTTTTGTATACCAAAAGAAATCTTTTAAAACTAGCCCTGATTTTAA	540	
QY	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200	
DB	541	GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGA	600	
QY	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220	
DB	601	GAATAAACTTACAAAAATAAACTATATGCTTATGCTTATGCTTATGCTTATGCTTAT	660	
QY	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr	240	
DB	661	ATAGACTAGCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAATAAATACTACACT	720	
QY	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260	
DB	721	AAAGATATCTCATACCTTCTAGTGAGCTTTATGGAATAATTTTCAAAAAATTTAATTTT	780	
QY	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280	
DB	781	AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	840	
QY	281	ProSerAspIleLysHisTyrPheLysGlyLys	291	
DB	841	CCTAGTGNATAAGCAATTTTCAAGAAAA	873	
RESULT 7				
US-10-735-419-8				
; Sequence 8, Application US/10735419				
; Patent No. 7026147				

GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel  
 ; APPLICANT: Wakarchuk, Warren W.  
 ; APPLICANT: National Research Council of Canada  
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
 ; FILE REFERENCE: 019633-000111US  
 ; CURRENT APPLICATION NUMBER: US/10/735,419  
 ; CURRENT FILING DATE: 2003-12-11  
 ; PRIOR APPLICATION NUMBER: US/09/816,028A  
 ; PRIOR FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/118,213  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: US 09/495,406  
 ; PRIOR FILING DATE: 2000-01-31  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 876  
 ; TYPE: DNA  
 ; ORGANISM: Campylobacter jejuni  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(876)  
 ; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
 ; OTHER INFORMATION: (CstII) from C. jejuni O:19  
 ; US-10-735-419-8

Alignment Scores:

Pred. No.: 3.66e-173 Length: 876  
 Score: 1563.00 Matches: 291  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 5 Gaps: 0

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCGCATTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGATTCTTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATTAAGAACTATATGCTTATGCTCTAATAGCTTTTACGAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATGAACATAGCGCCAAATTTAAATTTCAAAATTTTATCATCAAGAAAAATAACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAGATATATCTATACCTCTAGTCAGCTTATGCAAGAAATTTTCAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873

## RESULT 8

US-09-816-028A-13  
; Sequence 13, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36

## US-09-816-028A-13

Alignment Scores:  
Pred. No.: 1,07e-172 Length: 873  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
DB: 3 Gaps: 0

## US-10-734-719-9 (1-291) x US-09-816-028A-13 (1-873)

QY 1 MetLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATGCTGGAATGGCAAGTTTAAAGAAATTTGATTATTCAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTTCATGTTAGATGTAATCAATTTTATTTTTGNAGATAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAATATGAGACCGAACTAAATTTATGTGTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGlnAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTAATTTTCTCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTATTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTCAAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGTATCCCAACAAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATTAAGAACTATATGCTTATGCTCTAAAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATGAACATAGCGCCAAATTTAAATTTCAAAATTTTATCATCAAGAAAAATAACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAGATATATCTATACCTCTTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873

## RESULT 9

US-10-303-162-13  
; Sequence 13, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-162-13

Alignment Scores:
Pred. No.: 1.07e-172 Length: 873
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservatives: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-162-13 (1-873)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGCGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAAACAGTGTTCACCCCTAAATTTCTTTGAGCAATACATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrThrGluLeuLeuMetCysSerAsn 80
DB 181 ACTTTAAACATTTAAATCCAAAATCAAGATATGAGACCACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTACGATTAATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTACGAAAATTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTGCGAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGCGGAATGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTGATACCAACAAAGAAATCTTTTAAACACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATCTTCAAAATAAACTATTTGCTTATGCTCTTAATAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGACTAGCGCCCAAAATTTAAATTTCAAAATTTATCATACAGAAAAATAACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACCTACACTTCTAGTGAGGCTTATGGAAAATTTTCAAAAAATATTAATTTT 780

```

```

QY 261 LysLysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAAGCATTATTTCAAAGAAAA 873

RESULT 10
US-10-303-134-13
; Sequence 13, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:36
US-10-303-134-13

Alignment Scores:
Pred. No.: 1.07e-172 Length: 873
Score: 1559.00 Matches: 290
Percent Similarity: 99.7% Conservatives: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-134-13 (1-873)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGCGAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGATGATTTAGATGATTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAAACAGTGTTCACCCCTAAATTTCTTTGAGCAATACATAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrThrGluLeuLeuMetCysSerAsn 80
DB 181 ACTTTAAACATTTAAATCCAAAATCAAGATATGAGACCACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTACGATTAATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360

```

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTCCAGAAATTTATTTCAATCAAGAAATACCTCAGGGGTCTATATATGTGCGAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATGCTCACACTATATCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTACAAATATAAATTTTATGCTTAATGCTTAATAGCTTTTGTAGCAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
DB 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAAAGAAAAATAAATACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATATCTCATACCTCTTAGTGAGGCTTTATGGAATAATTTTCAAAATAATTAATTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAATAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

## RESULT 11

US-10-303-118-13  
; Sequence 13, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 873  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:36  
US-10-303-118-13

## Alignment Scores:

Pred. No.: 1,07e-172 Length: 873  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0

Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-118-13 (1-873)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAAGTTATTTATTGCTGGAATGACCAAGTTTAAAGAAATTCGATTATTTCAAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAATGATTTTGTATTTAGATGATCAATTTTATTTTGAAGATAATACTAT 120  
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTGTGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTGAGCAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATAGACCGAACTAATTAATGTGTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTTCC 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCAATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTTACGAAATTTATTTCAATCAAGAAATTTTAAACAACTTTTAAAGAAATTTAATGCTTATTTAAA 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATGCTCACACTATATCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTACAAATAAATAAATTTTAAATTTTATGCTTAATGCTTAATAGCTTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
DB 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATCATACAAAGAAAAATAAATACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATATCTCATACCTCTTAGTGAGGCTTTATGGAATAATTTTCAAAATAATTAATTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAATAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

## RESULT 12

US-10-303-128-13  
; Sequence 13, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.

	Qy	1	MetLysValleileAlaGlyAsnGlyProSerLeuLysGlulleAspTyrSerArg	20
	Db	1	ATGAATAAAGTTATTATCTCGAAATGGACCAAGTTTAAAAGAAMATTGATTATTCAGG	60
	Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTy+PheGluAspLysTyrTyr	40
	Db	61	CTRCCAAATGATTTTGATGTATTTAGATGTATCAATTTTATTTGAAAGATAAATACTAT	120
	Qy	41	LeuGlyLysLysCysLysAlaValPheTyThrProAsnPhePheGluGlnTyrTyr	60
	Db	121	CITGGTAAAAATGCACAAACAGTGTTTTACACCCTTAATTTCTCTTTGAGCAATACTAC	180
	Qy	61	ThrLeuLysHisLeulleGlnAsnGlnGluTyThrGluLeulleMetCysSerAsn	80
	Db	181	ACITTTAAACATTTAATCCAAAATCAAGAAATATGAGCCGAACATAATTATGTGTTCTAAT	240
	Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyRaspTyrPhePro	100
	Db	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCT	300
	Qy	101	AspAlaHisLeuClyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
	Db	301	GATGCTCATTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA	360
	Qy	121	PheHisglulleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
	Db	361	TTTCAGCAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTAGCC	420
	Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
	Db	421	ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTGATTTTTTATCAAAATGGGTCA	480
	Qy	161	SerTyxAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180
	Db	481	TCITTATGCTTTTGATACCAACCAAGAAAAATCTTTTAAAACTAGCCCCCTGATTTTAAAAAT	540
	Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200

Alignment Scores:		
Pred. No.:	1.07e-172	Length: 873
Score:	1559.00	Matches: 290
Percent Similarity:	99.7%	Conservative: 0
Best Local Similarity:	99.7%	Mismatches: 1
Query Match:	99.7%	Indels: 0
DB:	5	Gaps: 0
US-10-734-719-9 (1-291)	x	US-10-735-419-13 (1-873)

Qy	1	MetLysLysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg
Db	1	ATGAATAAAGTTATTATTGCTCGAAATGGACCAAGTTAAAAAGAAATTGATTATTCAAAG
Qy	21	LeuProAsnAspPheAspValIPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr
Db	61	CTACCAATAGATTTTTTGATGTATTTAGATGTAATCAATTTTATTTCGAAGATAAATACTAT
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr



```
Db 121 CTTGGTAAAAAATGCAAAACAGTGTGTTTACACCCCTAAATTTCTTCTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGAAGAACTAATATATGTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATGTGTGCAATAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATGCTCACACTATATCGACATAGTAAAAATACAGATATAAAAAGCTTTTAAAGTTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTCAAAAATTAATACTATATGCTTATGCTCTTAATAGCTTTTACCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATTAAGCAATTTTTCAAAGAAAA 873
```

## RESULT 14

```
US-09-816-028A-11
; Sequence 11, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
```

```
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3(alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (C8II) from C. jejuni O:4
US-09-816-028A-11
```

```
Alignment Scores:
Pred. No.: 1 83e-172 Length: 873
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-9 (1-291) x US-09-816-028A-11 (1-873)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTTATTTATTTCTGGAATGGACCAAGTTTAAAGAAATTTGATTATCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAAAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAATGCAAGACGCTGTTTACACCCCTGGTTCCTTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTTATCCAAATCAAGATATGACCGAAGCTAATATATGTGTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGTGCAATAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTGGGAAATTTGATTTTATCAAAAATGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATGCTCACACTATATCGACATAGTAAAAATACAGATATAAAAAGCTTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTCAAAAATTAATACTATATGCTTATGCTCTTAACAGCTTTTACCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTTAAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
```

QY	281	ProSerAspIleHisTyrPheLysGlyLys 291	Db	361	TTTCACGAAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTGTGCGTAGCC 420
Db	841	CCTAGTGATATAAGCAATTTATTTCAAAGGAAAA 873	QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGluAsnGlySer 160
			Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
			QY	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
			Db	481	TCATTATGCTTTTGTATACCAACAGAAAAATCTTTTAAACTAGCCCTGATTTTAAAT 540
			QY	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
			Db	541	GATCGCTCACATATATCGGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTTCTA 600
			QY	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
			Db	601	GAATAAACTTACAAAATAAAACTATATTTGCTTATGTCTCTAAACAGTCTTTTAGCAATTTT 660
			QY	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
			Db	661	ATAGAACTAGCGCAAAATTTTAAATTCAAATTTTATCATACAAGAAAAATAACTACT 720
			QY	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
			Db	721	AAAGATATACTATACCTTCTAGTGAGCTTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
			QY	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
			Db	781	AAAAAAATAAAATAAAAGAAATGTTTATTATACAGTTGTATAAAGATCTATTAAAGATTA 840
			QY	281	ProSerAspIleLysHisTyrPheLysGlyLys 291
			Db	841	CCTAGTGATATAAGCAATTTATTTCAAAGGAAAA 873
					RESULT 16
					US-10-303-134-11
					; Sequence 11, Application US/10303134
					; Patent No. 6825019
					; GENERAL INFORMATION:
					; APPLICANT: Gilbert, Michel
					; APPLICANT: Wakarchuk, Warren W.
					; APPLICANT: National Research Council of Canada
					; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
					; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
					; FILE REFERENCE: 019633-000111US
					; CURRENT APPLICATION NUMBER: US/10/303,134
					; CURRENT FILING DATE: 2002-11-21
					; PRIOR FILING DATE: 2001-03-21
					; PRIOR FILING DATE: 2001-03-21
					; PRIOR APPLICATION NUMBER: US 60/118,213
					; PRIOR FILING DATE: 1999-02-01
					; PRIOR APPLICATION NUMBER: US 09/495,406
					; PRIOR FILING DATE: 2000-01-31
					; NUMBER OF SEQ ID NOS: 49
					; SOFTWARE: PatentIn Ver. 2.1
					; SEQ ID NO 11
					; LENGTH: 873
					; TYPE: DNA
					; ORGANISM: Campylobacter jejuni
					; FEATURE:
					; NAME/KEY: CDS
					; LOCATION: (1)..(873)
					; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
					; OTHER INFORMATION: (CstII) from C. jejuni O:4
					US-10-303-134-11
					Alignment Scores:
					Pred. No.: 1,83e-172 Length: 873
					Score: 1557.00 Matches: 290
					Percent Similarity: 99.7% Conservatives: 0
					Best Local Similarity: 99.7% Mismatches: 1
					Query Match: 99.6% Indels: 0
					DB: 3 Gaps: 0
					US-10-734-719-9 (1-291) x US-10-303-162-11 (1-873)
QY	1	MetLysLysValIleLeAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20	QY	1	MetLysLysValIleLeAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db	1	ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60	Db	1	ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG 60
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40	QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db	61	CTACCAATGATTTGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120	Db	61	CTACCAATGATTTGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
QY	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60	QY	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
Db	121	CTTGGTAAAAATCAAGACAGTGTTTTACACCTCGTTCCTTTTGACCAATACTAC 180	Db	121	CTTGGTAAAAATCAAGACAGTGTTTTACACCTCGTTCCTTTTGACCAATACTAC 180
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80	QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db	181	ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATATATGTGTTCTAAT 240	Db	181	ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATATATGTGTTCTAAT 240
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100	QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300	Db	241	TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCCT 300
QY	101	AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120	QY	101	AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360	Db	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140	QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

```
US-10-734-719-9 (1-291) x US-10-303-134-11 (1-873)
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni O:4
US-10-303-118-11

Alignment Scores:
Pred. No.: 1.83e-172 Length: 873
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-118-11 (1-873)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTCGTAATAAATGCAAGACAGTGTTCACCCCTGGTTCTTCTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAATTAATGTTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGlnPheTyrPheTyrAspTyrPhePro 100
Db 241 TACCAACAGCTCATCTAGAAATGAAATTTTGAAAACTTTTACGATTAATTTCTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACCTTAAAGAAATTAATGCTTTATTA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAAGAAATTTTCAATCAAAAGATTAACCTCAGGGGTCTATATGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAAGATTT 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATTGCTTATGCTCTAACAGTCTTTTGAAGAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAAATACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAGATATACTATACATCTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATATTAATTT 780
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATAAAGCATTAATTTTCAAAAGGAAAA 873

RESULT 17
US-10-303-118-11
; Sequence 11, Application US/10303118
; Patent No. 690867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
```

```
Db 601 GAAAAAATTACAAAATAAAATATATGCTTATGCTTAAACAGTCTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPhelileGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTCAAATTTATCATACAAGAAAAATAACTACACT 720
QY 241 LysAspileuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATAAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspileLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTAATTTCAAGGAAAA 873

RESULT 18
US-10-734-128-11
; Sequence 11, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(873)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
US-10-734-128-11
Alignment Scores:
Pred. No.: 1.83e-172 Length: 873
Score: 1557.00 Matches: 290
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 1
Query Match: 99.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-128-11 (1-873)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGGTGTATTTAGATGTAATCAATTTTATTTGAAGATAAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAATGCAAGCAGTGTTTTACACCCCTGTTCTTCTTTGAGCAATACTAC 180
```

```
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACITTTAAACATTTTATCTCAAAATCAAGATATGAGACCGAAGCTAATATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGACGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGCATATGTAATAATACAGATATATAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTCTAAACAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPhelileGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720
QY 241 LysAspileuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATAAAGAAATGTTTATTAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspileLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 19
US-10-735-419-11
; Sequence 11, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
```

FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(873)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (CstII) from C. jejuni O:4  
US-10-735-419-11

Alignment Scores:  
Pred. No.: 1.83e-172 Length: 873  
Score: 1557.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.6% Indels: 0  
DB: 5 Gaps: 0

US-10-734-719-9 (1-291) x US-10-735-419-11 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60  
Db 121 CTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCTTCTTTGAGCAATAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80  
Db 181 ACTTTAAACATTAAATCCAAATCAAGAAATATGAGCCGAACTAAATATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGTAATTTGTAAAAACTTTTACGATTAATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACCTTAAAGAAATTAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGATGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATACAAAAATAAACTATATGCTTATGCTCTTAACAGCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTATCATCAAGAAAAATAAATACTACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATATAAAATAAGAAATATGTTTATACAAATTTGATAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAAGCATTTATTTTCAAGGAAAA 873

RESULT 20  
US-09-495-406-4  
; Sequence 4, Application US/09495406  
; Patent No. 6503744  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; TITLE OF INVENTION: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000110US  
; CURRENT APPLICATION NUMBER: US/09/495,406  
; CURRENT FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of LOS biosynthesis locus)  
US-09-495-406-4

Alignment Scores:  
Pred. No.: 4.52e-169 Length: 876  
Score: 1528.00 Matches: 285  
Percent Similarity: 98.3% Conservative: 1  
Best Local Similarity: 97.9% Mismatches: 5  
Query Match: 97.8% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-09-495-406-4 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTGATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60  
Db 121 CTGGTAAAAAATTTCAAGCAGTATTTTACAAATCTGGGCTTTTTCGAACAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80  
Db 181 ACTTTAAACATTAAATCCAAATCAAGAAATATGAGCCGAACTAAATATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAAACTTTTACGATTAATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATGATGATTTTAAACAACCTTAAAGAAATTAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTTGCGATGCT 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTGATACCAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATACAAAAATAAACTATATGCTTATGCTCTTAACAGCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTATCATCAAGAAAAATAAATACTACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATATAAAATAAGAAATATGTTTATACAAATTTGATAAAGATCTATTAAAGATTA 840

```
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAAATGATTTTATCAAAATGGCTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTAAAACTGGCTCCTCATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAAATATATGCTTATGCTTAAACAGCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATCAAGAAAAATAAATACTACTACT 720
Qy 241 LysAspIleLeuLysProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATCACTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTTAAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATATTTATCAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 22
US-09-816-028A-4
; Sequence 4, Application US/09816028A
; Patent No. 6693705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-11-18,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-09-816-028A-4

Alignment Scores:
Pred. No.: 4,52e-169 Length: 876
Score: 1528.00 Matches: 285
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-09-816-028A-4 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg 20
```

```
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTGATGTATTAGTAGCAATCAATTTTATTGAAAGATAAATACTAT 120
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAATAATTCAGACGATTTTACATCCTGGCTCTTTTGTGAAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTTATCCAAATCAAGATATAGAGCCGAACCTAATATGTGTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTTCAGATTTATTTCC 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTTAATGCTTATTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATCTCAATCAAGAAATTTACCTCAGAGTCTATATGTGTGCAAGTAG 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAAATTTGATTTTATCAAAATGGG 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAAAATCTTTTAAAACTGGCTCCTCATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAACTTTACAAAAATAAAACTATATTTGCTTATGCTTCTACAGTCTTTTAGCAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTTAAATTCAAATTTTATCATCAAGAAAAATAAATACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTATCACTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAAAATATTTATTAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTTCAAGGAAAA 873

RESULT 22
US-10-303-162-4
; Sequence 4, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
```

```
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-4
Alignment Scores:
Pred. No.: 4.52e-169 Length: 876
Score: 1528.00 Matches: 285
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 3 Gaps: 0
US-10-734-719-9 (1-291) x US-10-303-162-4 (1-876)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGGAATATGCAAGTAAAGAAATGATTTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTGATGATTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTCGTAATAAATTCAAAGCAGTATTTTACAATCTGCTCTTTTGTGAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAACTAATATGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATGATGATTTTAAACCACTTAAAGAAATTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTCAAGAAATTTATCTCAATCAAGAATTAACCTCAGGAGTCTATATGTCGAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTGATACCAACCAAGAAATCTTTTAAACCTGGCTCTCGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAACCTTACAAAATPAAACTATATTCGTTATGTCCTAACAGTCTTTTACGAAATTTT 660
```

```
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTATCATACAAGAAAAATACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAGATATATCTCATCTCTAGTGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAAAAATTAAGAAAAATATTTATTAACAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATTAAGCATTTATTTCAAGGAAAA 873
```

## RESULT 23

```
US-10-303-134-4
; Sequence 4, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-134-4
```

```
Alignment Scores:
Pred. No.: 4.52e-169 Length: 876
Score: 1528.00 Matches: 285
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 3 Gaps: 0
US-10-734-719-9 (1-291) x US-10-303-134-4 (1-876)
```

```
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTCGGAATATGCAAGTAAAGAAATGATTTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTGATGATTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTCGTAATAAATTCAAAGCAGTATTTTACAATCTGCTCTTTTGTGAACAATACTAC 180
```

FEATURE:		NAME/KEY: CDS	
LOCATION: (1)..(876)		OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase	
OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni		OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)	
OTHER INFORMATION: biosynthesis locus)		US-10-303-118-4	
Alignment Scores:		4.52e-169	Length: 876
Pred. No.:		1528.00	Matches: 285
Percent Similarity:		98.3%	Conservative: 1
Best Local Similarity:		97.9%	Mismatches: 5
Query Match:		97.8%	Indels: 0
DB:		3	Gaps: 0
US-10-734-719-9 (1-291) x US-10-303-118-4 (1-876)			
QY	1	MeLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGG	60
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr	40
Db	61	CTACCAATGATTTTGTATGATTTAGATGCATCAATTTTATTATTGAAGATAAATACTAT	120
QY	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr	60
Db	121	CTTGGTAAAAATTTCAAGCAGTATTTTACAATCTCGTCTTTTGTGAAACAATACTAC	180
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAACATTTTAAATCCAAATCAAGAAATATGACGCGAACTAATTTATGTTCTAAT	240
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTATTTCCT	300
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATGATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTAAAA	360
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTATCTCAATCAAGAAATTTATCTTCGGAATTTGATTTTCAAAATGGTCA	420
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAATTTGATTTTCAAAATGGTCA	480
QY	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCCTATGCTTTGATACCAACAAAGAAATCTTTTAAAACTGGCTCTGATTAAAAAAT	540
QY	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA	600
QY	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAACTTACAAATAAAACTATATTGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT	660
QY	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACTACT	720
QY	241	LysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATACATCACTCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTT	780
QY	261	LysLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	841	CCTAGTGATATAAGCAATTTTCAAGGAAAA 873	
RESULT 24			
US-10-303-118-4			
Sequence 4, Application US/10303118			
Patent No. 6905867			
GENERAL INFORMATION:			
APPLICANT: Gilbert, Michel			
APPLICANT: Wakarchuk, Warren W.			
APPLICANT: National Research Council of Canada			
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of			
TITLE OF INVENTION: Gangliosides and Ganglioside Mimics			
FILE REFERENCE: 019633-000111US			
CURRENT APPLICATION NUMBER: US/10/303,118			
CURRENT FILING DATE: 2002-11-21			
PRIOR FILING DATE: US/09/816,028			
PRIOR FILING DATE: 2001-03-21			
PRIOR FILING DATE: US 60/118,213			
PRIOR FILING DATE: 1999-02-01			
PRIOR FILING DATE: US 09/495,406			
PRIOR FILING DATE: 2000-01-31			
NUMBER OF SEQ ID NOS: 49			
SOFTWARE: Patentin Ver. 2.1			
SEQ ID NO 4			
LENGTH: 876			
TYPE: DNA			
ORGANISM: Campylobacter jejuni			



Db 781 AAAAAATAAAATTAAGAAAAATATTTATACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTATATAAAGCAATTTTCAAAAGGAAAA 873

RESULT 25  
US-10-734-128-4  
; Sequence 4, Application US/10303128  
; Patent No. 6911337  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,128  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-303-128-4

Alignment Scores:  
Pred. No.: 4.52e-169 Length: 876  
Score: 1528.00 Matches: 285  
Percent Similarity: 98.3% Conservative: 1  
Best Local Similarity: 97.9% Mismatches: 5  
Query Match: 97.8% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-128-4 (1-876)

Qy 1 MetLysLysValIleLeuAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATGCTGGAATGACCAAGTTAAAGAAATTTGATTATTTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTGTATGATTTAGTATGCAATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGGTAAATAAATTCAAAGCAGATTTTACAAATCCTGGTCTTTTGTGAACAATCTACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAATCAAGATATGACCGCAACTAATTTATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCCCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATGATGATTTTTTTTAAACAACTTAAAGAAATTTAATGCTTTATTTAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCCAGCAATTTATCTCAATCAAGAAATTAACCTCAGGAGTCTATATGTGTGCTAGTCT 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyLysLeuAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTGGCTCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGCACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAATAAAACTATATTTGCTTATGCTCTAACAGTCTTTTAGCAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTTCTAGTGAGCTTATGGAATTTTCAAAAAATTTAATTTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAAATTAAGAAAAATATTTATACAAAGTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTATATAAAGCAATTTTCAAAAGGAAAA 873

RESULT 26  
US-10-735-419-4  
; Sequence 4, Application US/10735419  
; Patent No. 7026147  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/735,419  
; CURRENT FILING DATE: 2003-12-11  
; PRIOR APPLICATION NUMBER: US/09/816,028A  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase  
; OTHER INFORMATION: Campylobacter sialyltransferase II (csII) from C. jejuni  
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)  
; OTHER INFORMATION: biosynthesis locus)  
US-10-735-419-4

Alignment Scores:  
Pred. No.: 4.52e-169 Length: 876  
Score: 1528.00 Matches: 285  
Percent Similarity: 98.3% Conservative: 1  
Best Local Similarity: 97.9% Mismatches: 5  
Query Match: 97.8% Indels: 0  
DB: 3 Gaps: 0

```

Score: 1528.00 Matches: 285
Percent Similarity: 98.3% Conservativity: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 5 Gaps: 0

US-10-734-719-9 (1-291) x US-10-735-419-4 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTTCCTCGAAATGGACCAAGTTTAAAGAAATGATTGATTCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGAATGATTAGATCAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATTCAAAGCAGTATTTTACAATCCTGGTCTTTTGTGAAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCCGAACATAATTATGTCTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCAATTCGGATATGATTTTAAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCCAGAAATTTATCTCAATCAAGAAATTTACCTTCAGAGTCTATATGTCGAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAsnPheLysAsn 180
DB 481 TCTTATGCTTTTGATCAACCAAGAAATCTTTTAAAACTGGCTCTCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
QY 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAATATAAATCTATATGCTTATGCTCTTACAGCTTTTGTAGCAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAATTTTAAATTTCAATTTTATCATACAAGAAAAATTAATACACT 720
QY 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTCTTAGTGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
QY 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAATAAAMTTAAAGAAAAATTTATTTTACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTATATAAGCATTTTTCAAAGGAAAA 873

RESULT 27
US-09-495-406-2
; Sequence 2, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of LOS biosynthesis locus)
US-09-495-406-2

Alignment Scores: 1,01e-168 Length: 876
Pred. No.: 1525.00 Matches: 283
Score: 98.6% Conservativity: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-09-495-406-2 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTTCCTGGAAATGGACCAAGTTTAAAGAAATGATTGATTCAAGA 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGAATGATTAGATCAATCAATTTTATTTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTTGGTAAAAATGCAAGGCAGTATTTTACAATCTCTATCTTTTGTGAAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACATAATTATGTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTTCTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCAATTCGGATATGATTTTTCACCAACCTTAAAGAAATTTAATGCTTATTTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTCCAGAAATTTATCTCAATCAAGAAATTTACCTTCAGAGTCTATATGTCGAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTCATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAsnPheLysAsn 180
DB 481 TCTTATGCTTTTGATCAACCAAGAAATCTTTTAAAACTGGCTCTCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
QY 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACAAATATAAATCTATATGCTTATGCTCTTACAGCTTTTGTAGCAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
DB 661 ATAGAACTAGCGCAATTTTAAATTTCAATTTTATCATACAAGAAAAATTAATACACT 720
QY 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACTCATACCTCTTAGTGGCTTATGGAATTTTCAAAAAATTTAATTTT 780
QY 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAATAAAMTTAAAGAAAAATTTATTTTACAAGTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTATATAAGCATTTTTCAAAGGAAAA 873

```



[illegible]

Patent No. 6825019  
GENERAL INFORMATION:  
APPLICANT: Gilbert, Michel  
APPLICANT: Wakarchuk, Warren W.  
APPLICANT: National Research Council of Canada  
TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
Gangliosides and Ganglioside Mimics

; CURRENT APPLICATION NUMBER: US/10/303,134

```

; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3(alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jej
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-303-134-2

Alignment Scores:
Pred. No.: 1.01e-168 Length: 876
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-134-2 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTCGTGGAATGACCAAGTTTAAAGAAATTCGATTATTCAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGTGATGATTATAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGGCAGTATTTTACAATCCTTATTCCTTTTTTTTGAAACAATAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAAAACATTTTAATCCAAAATCAAGAAATATGAGCCGAACCTAATTAATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCCCT 300

```

US-10-734-719-9 (1-291) x US-10-303-134-2 (1-876)

Oy		1	MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGlulleaspTyrSerArg	20
Db		1	ATGAATAAAGATTATTTCCTGGAAATGACCACCAAGTTAAAAAAAATTGAATTAACAAGA	60
Oy		21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyPheGluAspLysTyrTyr	40
Db		61	CTACCAAATGATTTTTTGATGTATTTAGATGCTAACATTTTTTTTTCGAAGATAAATACTAT	120
Oy		41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db		121	CTTGCGTAAAAAATCCAAGSCAGTATTTTACAATCCCTATTCTTTTTTTTGAAACAATACTAC	180
Oy		61	ThrLeuLysHisLeuLecIhnAsnGlnGluTyrGluThrGluLeuilleMetCysSerAsn	80
Db		181	ACTTTTAAAAACATTTTAATCCAAATCAAGAATATGAGACCGAAGCTAATTATGTGTCTAAT	240
Oy		81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhPro	100
Db		241	TACAACCAAGCTCATCTACAAAAATGAAAAATTTTGTAAAAACTTTTTTACGATTAATTTTCCT	300
Oy				120

```
Db 301 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTATTTTCAATCAAGAATAATACCTCAGGGGTTTATATGTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACACAAAGAAATTTAATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACATAACAAACAAAATCTTTTAAATTTGGTCTCTTAATTTTAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTAAATTTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAATAAATACTATATTTGCTTATGTCCTAACAGTCTTTTACGAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGTCAGGCTTATGAAAAATTTTCAAAAAATATTTAAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAATAAATAAAGAAAAATTTTATACAAAGTTGATAAAGATCTATTAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 31
US-10-303-118-2
; Sequence 2, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-118-2
```

```
Alignment Scores:
Pred. No.: 1,01e-168 Length: 876
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-118-2 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGGTTATTTATTCGTCGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAAAATGATTTTGTATGATTTATGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGCAGTATTTTACAATCCTATTTCTTTTGTGACCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTGTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAGGCTCATCTAGAAAATGAAAAATTTTGTAAAAAATTTTACCGATTATTTTCCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCCAGAAATTTTATTTCAATCAAGAAATTTTACCTCAGGGGTTTATATGTGTCAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACATAAACAATAAATCTTTTAAAAATTTGGCTCTCTAATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGCATAGTAAATAACAGATATAAAGCTTTTAAATTTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACAAAATAAATACTATATTTGCTTATGCTTATGCTTAAACAGTCTTTTAGCAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTTAAATTTCAATTTTATCATACAAGAAAAATAACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAATAAATAAAGAAAAATTTTATTAACAAGTTTGAATAAAGATCTATTAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 32
US-10-303-128-2
; Sequence 2, Application US/10303128
```

Patent No. 6911337		161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
; GENERAL INFORMATION:		481	TCATTATGCTTTTGGATCTAAACAAATAATCTTTTAAAAATGGCTCCTCTTAATTTTAAAAAT	540
; APPLICANT: Wakarchuk, Warren W.		181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
; APPLICANT: National Research Council of Canada		541	GATAATTCACACTATATCGCACATAGTAAATAACACAGATATAAAAAAGCTTTAGAATTTCTA	600
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of		201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuLeuAsnPhe	220
; FILE REFERENCE: 019633-000111US		601	GAATAAACTTACAAATAAAACTATATTGCTTATGTCTTACAGTCTTTTAGCAAAATTTT	660
; CURRENT APPLICATION NUMBER: US/10/303,128		221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
; CURRENT FILING DATE: 2002-11-21		661	ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATCAAGAAAAATAACTACTACT	720
; PRIOR FILING DATE: 2001-03-21		241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
; PRIOR APPLICATION NUMBER: US 60/118,213		721	AAAGATATACTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAAATTTT	780
; PRIOR FILING DATE: 1999-02-01		261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
; PRIOR APPLICATION NUMBER: US 09/495,406		781	AAAAAATAAAAAATTAAGAAAAATATTATTACAAAGTTGATAAAAAAGATCTATTAAAGATT	840
; PRIOR FILING DATE: 2000-01-31		281	ProSerAspIleLysHisTyrPheLysGlyLys	291
; NUMBER OF SEQ ID NOS: 49		841	CCTAGTGATATAAAGCATTTATTTCAAAGGAAAA	873
; SOFTWARE: Patentin Ver. 2.1		RESULT 33		
; SEQ ID NO 2		US-10-735-419-2		
; LENGTH: 876		; Sequence 2, Application US/10735419		
; TYPE: DNA		; GENERAL INFORMATION:		
; ORGANISM: Campylobacter jejuni		; APPLICANT: Wakarchuk, Warren W.		
; NAME/KEY: CDS		; APPLICANT: National Research Council of Canada		
; LOCATION: (1)..(876)		; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of		
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase		; FILE REFERENCE: 019633-000111US		
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni		; CURRENT APPLICATION NUMBER: US/10/735,419		
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)		; CURRENT FILING DATE: 2003-12-11		
; OTHER INFORMATION: biosynthesis locus)		; PRIOR FILING DATE: 2001-03-21		
US-10-303-128-2		; PRIOR APPLICATION NUMBER: US 60/118,213		
Alignment Scores:		; PRIOR FILING DATE: 1999-02-01		
Pred. No.:	1,01e-168	Length:	876	
Score:	1525.00	Matches:	283	
Percent Similarity:	98.6%	Conservative:	4	
Best Local Similarity:	97.3%	Mismatches:	4	
Query Match:	97.6%	Indels:	0	
DB:		Gaps:	0	
US-10-734-719-9 (1-291) x US-10-303-128-2 (1-876)		US-10-734-719-9 (1-291) x US-10-303-128-2 (1-876)		
QY	1	MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg	20	
DB	1	ATGAAAAAGCTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGA	60	
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr	40	
DB	61	CTACCAATGATTTTGATGTATTAGATGTAATCAATTTTATTTTGAAGATAAAATCTAT	120	
QY	41	LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60	
DB	121	CTTGGTAAAAAATGCAAGGAGATATTTTACAACTCTTCTTTTGAACATACTAC	180	
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80	
DB	181	ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAAGCTAAATATGCTGTTCTAAT	240	
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValIysThrPheTyrAspTyrPhePro	100	
DB	241	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATATTATTTTCC	300	
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120	
DB	301	GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGATTTTAAATGCTTATTAAAA	360	
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140	
DB	361	TTTCACGAATTTATTTCAATCAAGAAATTAACCTCAGGGTTTATATGTCGAGTAGCC	420	
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160	
DB	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA	480	
QY		1 MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg		
DB				

```
Db 1 ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysCysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTCGTAATAAATGCAAGGCAGTATTTTACAATCTTCTTTTGTGAACTATTTTGAACATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAATTAATTTATGTTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTTACCTCAGGGTTTATATATGTCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTGATACCTAAACAAAATAATCTTTTAAATTTGGCTCTCTTAATTTTAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGCATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATATAAATTTTATGCTTATGCTTACAGCTTTTACGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTACGCGCAATTTTAAATTTTCAATTTTATCATACAAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTATACCTTTCTAGGAGCTTATGGAATAATTTTCAAAAAATTTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAATAAAGAAATATTTTATTTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTTCAAGGAAAA 873
```

## RESULT 34

```
US-09-495-406-1
; Sequence 1, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
```

```
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-09-495-406-1
```

## Alignment Scores:

```
Alignment Scores:
Pred. No.: 3,966-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-9 (1-291) x US-09-495-406-1 (1-11474)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6168 CTTCGTAATAAATGCAAGGCAGTATTTTACAATCTCTTTTGTGAACTATTTTGAACATACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAATTAATTTATGTTCTTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTTATTTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCGAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGTATCTAAACAAAATAATCTTTTAAATTTGGCTCTCTTAATTTTAAAAT 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAAATTCACACTATATCGGCATAGTAAATAATACAGATATAAAGCTTTTAGAATTTCTA 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTACAAAATATAAATCTATTTGCTTATGCTCTTACAGCTCTTTTAGCAATTTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCAATTTTAAATTTTCAATTTTATCATACAAAGAAAAATAACTACACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATCTATCATCTCTTCTAGGAGCTTATGGAATAATTTTCAAAAAATTTAATTTT 6827
```

Tue Aug 22 15:02:29 2006

```
QY 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysAspLeuArgLeu 280
Db 6828 AAAAAAAAAAATAAAGAAATATTTATCAAGTTGATAAAGATCTATTAAGATTA 6887
QY 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTATATAAAGCATTATTTCAAGGAAAA 6920

RESULT 35
US-09-816-028A-1
; Sequence 1, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-09-816-028A-1

Alignment Scores:
Pred. No.: 3 96e-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
Gaps: 3

US-10-734-719-9 (1-291) x US-09-816-028A-1 (1-11474)

QY 1 MetLysLysValLleLleAlaGlyAsnGlyProSerLeuLysGluLysAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATCAAGA 6107
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 6108 CTACCAAAATGATTGATGATTTAGATGTAATCAATTTTATTGAGATAAATACTAT 6167
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
Db 6168 CTGGTAAAAAATGCAAGGCGAGTATTTTACAACTCTATCTTTTGTGACATACTAC 6227
QY 61 ThrLeuLysHisLeuLleGlnAsnGlnGlyThrGluLysLeuLysMetCysSerAsn 80
Db 6228 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATATGTTCTAAT 6287
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCAGAAAAATGAAATTTGTAAACCTTTTACGATATTTTCCCT 6347
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATGATGATTTTCAACCAACTTAAAGATTTTAAATGCTTATTTAAA 6407
QY 121 PheHisGluLysTyrPheAsnGlnArgLysThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGTCTATATGTGTGAGTACCC 6467
QY 141 IleAlaLeuGlyTyrLysGluLysLeuSerGlyLysPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTC 6527
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATCTACTAAACAAAAATCTTTTAAATTTGGCTCTCTAATTTTAAAA 6587
QY 181 AspArgSerHisTyrLysGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCT 6647
QY 201 GluLysThrTyrLysLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTACAAAAATAAACTATATTTGCTTATGCTTCTAACAGTCTTTTAGCAATTT 6707
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheLleLleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCCAAATTTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 6767
QY 241 LysAspLysLeuLysProSerSerGluAlaTyrGlyLysPheSerLysAsnLysPhe 260
Db 6768 AAAGATATATCATACCTTCTAGTGAGGCTTATGCAAAATTTTCAAAAAATATTAATTT 6827
QY 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu 280
Db 6828 AAAAAAATAAAATTAAGAAAAATATTTATTTACAGTTGATAAAGATCTATTAAGATTA 6887
QY 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTATATAAAGCATTATTTCAAGGAAAA 6920

RESULT 36
US-10-303-162-1
; Sequence 1, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-162-1

Alignment Scores:
Pred. No.: 3 96e-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
Gaps: 3

US-10-734-719-9 (1-291) x US-10-303-162-1 (1-11474)
```



```
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 6048 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAATAATGATTATTCAGA 6107
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 6108 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTATTTTGAAGATAAATACAT 6167
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 6168 CTTGTGTAATAAATGCAAGCGAGTATTTTACAAATCCTATTCTTTTGTGAACAATAC 6227
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 6228 ACTTTAAACATTTTAAATCCAAATATCAAGATATGAGCCGAATTAATGTTCTAAT 6287
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnValLysThrPheTyrAspTyrPhePro 100
DB 6288 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 6347
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 6348 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAA 6407
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 6408 TTTCAACGAATTTTATTTCAATCAAGAAATTTTGTAAAACTTTTACGATTATTTTCT 6467
QY 141 IleAlaLeuGlyTyrLysGluIleTyrSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 6468 ATAGCCCTAGGATACAAAGAAATTTTCTTTCGGGAATGATTTTATCAAAATGGGTCA 6527
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 6528 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAAAATTTGGCTCTCTAAATTTTAAA 6587
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 6588 GATAAATTCACACTATATCGGCATAGTAAAAATACAGATATATAAAGCTTTTAAAGATTA 6647
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 6648 GAAAAAATCTTACAAAATAAATACTATATGCTTATGCTTACAGCTCTTTTAGCAAAATTTT 6707
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleLysGlnGluLysAsnTyrThr 240
DB 6708 ATAGACTAGGCCCAATTTTAAATTTCAATTTTATCATAACAAGAAAAATAACTACTACT 6767
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 6768 AAAGATATCTCATACACTTCTAGTGAGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6827
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 6828 AAAAAATAAATAAATAAAGAAATATTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 6887
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 6888 CTTAGTGATATAAGCAATTTTTCAAAGGAAAA 6920
```

## RESULT 37

```
US-10-303-134-1
; Sequence 1, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
```

```
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-303-134-1
```

## Alignment Scores:

```
Pred. No.: 3 96e-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0
```

US-10-734-719-9 (1-291) x US-10-303-134-1 (1-11474)

```
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 6048 ATGAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAGAATAATGATTATTCAGA 6107
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 6108 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTATTTTGAAGATAAATACTAT 6167
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 6168 CTTGTGTAATAAATGCAAGCGAGTATTTTACAAATCCTATTCTTTTGTGAACAATAC 6227
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 6228 ACTTTAAACATTTTAAATCCAAATATCAAGATATGAGCCGAATTAATGTTCTAAT 6287
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnValLysThrPheTyrAspTyrPhePro 100
DB 6288 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 6347
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 6348 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAA 6407
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 6408 TTTCAACGAATTTTATTTCAATCAAGAAATTTTGTAAAACTTTTACGATTATTTTCT 6467
QY 141 IleAlaLeuGlyTyrLysGluIleTyrSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 6468 ATAGCCCTAGGATACAAAGAAATTTTCTTTCGGGAATTTTATGCTTATTTTCTAAT 6527
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 6528 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAAAATTTGGCTCTCTAAATTTTAAA 6587
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 6588 GATAAATTCACACTATATCGGCATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA 6647
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 6648 GAAAAAATCTTACAAAATAAATACTATATGCTTATGCTTACAGCTCTTTTAGCAAAATTTT 6707
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleLysGlnGluLysAsnTyrThr 240
```

```
Db 6708 ATAGAACTAGCCCAAAATTTAAATTTTATCATCAAGAAAAAATTAACACT 6767
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATATCTCATACCTTCTAGTGGGCTTATGGAATTTTCAAAAAATATTAAATTT 6827
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 6828 AAAAAAATAAAATTAAGAAAAATTTTATCAAGTTGATAAAAGATCTATTAAAGATTA 6887
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCAATTTATTCAAGGAAAA 6920

RESULT 39
US-10-303-118-1
; Sequence 1, Application US/10303118
; Patent No. 6905867
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/10/303,118
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
US-10-303-118-1
Alignment Scores:
Pred. No.: 3.96e-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-118-1 (1-11474)
QY 1 MetLysLysValIleLeuAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATCAAGA 6107
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 6108 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 6167
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
Db 6168 CTTGGTAAAAAATGCAAGGCGAGTATTTTACAATCCTATTTCTTTTGAACAATACTAC 6227
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 6228 ACTTTAAACAATTTTAAATCCAAATCAAGAATATGAGACCGAAGTAATATGTTCTTAAT 6287
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
```

```
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTTCGATTATTTTCTCT 6347
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATGATGATTTTTCAAACAACTTAAAGATTTTAAATGCTTATTTTAAA 6407
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTGTGACGTAGCC 6467
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 6527
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACTAAACAAAAAATCTTTTAAAAATTTGGCTCCTTAATTTTAAAAAT 6587
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6647
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTACAAAAATAAACTATATTGCTTATGCTCTAACAGTCTTTTAGCAAAATTT 6707
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCCAAATTTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTACACT 6767
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATATCTCATACCTTCTAGTGGGCTTATGGAATTTTCAAAAAATATTAAATTT 6827
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAAATAAAATTTAAAGAAAAATATTATTACAAGTTTGATAAAAGATCTATTAAAGATTA 6887
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCAATTTATTCAAGGAAAA 6920

RESULT 39
US-10-303-128-1
; Sequence 1, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
US-10-303-128-1
Alignment Scores:
```

```
Pred. No.: 3.96e-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-128-1 (1-11474)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATGCTGGAAATGGCAAGTTTAAAGAAATGATTATTCAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 6108 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAATAC 6167
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6168 CTTGGTAAAAATGCAAGGCAGTATTTTACAATCTTATCTTTTGAACATACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAATTAATATGTTCTA 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTCCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGATGATGATTTTCAACAACCTTAAAGATTTTATGCTTTTAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGCGGTCTATATGTCAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATCTAAACAAAGAAATCTTTTAAAAATGGCTCTCTAATTTTAA 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 6647
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTACAAAATAAAATATATTTGCTTATGCTTACAGCTCTTTTACGAAATTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 6708 ATGAAGTACGGCAATTTTAAATTTCAATTTTATCATCAAGAAAAATACTACACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATACATCATCTCTAGTGGCTTATGGAAATTTTCAAAAAATATTAATTTT 6827
Qy 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAATAAAATTAAGAAAAATATTTATCAAGTTTGATAAAGATCTATTAAAGATTA 6887
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTATATAAGCAATTAATTTCAAGGAAAA 6920
```

RESULT 40

US-10-735-419-1

; Sequence 1, Application US/10735419

; Patent No. 7026147

```
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735.419
; CURRENT FILING DATE: 2003-12-11
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-10-735-419-1
```

```
Alignment Scores:
Pred. No.: 3.96e-167 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 5 Gaps: 0

US-10-734-719-9 (1-291) x US-10-735-419-1 (1-11474)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATGCTGGAAATGGCAAGTTTAAAGAAATGATTATTCAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAATAC 6167
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6168 CTTGGTAAAAATGCAAGGCAGTATTTTACAATCTTATCTTTTGAACATACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAATTAATATGTTCTA 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTCCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGATGATGATTTTCAACAACCTTAAAGATTTTATGCTTTTAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCCAGAAATTTATTTCAATCAAGAAATTAACCTCAGCGGTCTATATGTCAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATCTAAACAAAGAAATCTTTTAAAAATGGCTCTCTAATTTTAA 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
```

Db	6588	GATAATTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA	6647
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	6648	GAAAAAAGTTACAAAAATAAACTATATGCTTATGCTTACAGTCTTTTACGCAATTTT	6707
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	6708	ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTACACT	6767
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	6768	AAAGATATATCTACATCTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT	6827
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	6828	AAAAAAATAAAAATTAAGAAAAATATTTATCAAGTTGATAAAGATCTATTAAAGATT	6887
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	6888	CCTAGTGATATAAAGCAATTATTTCAAGGAAAA	6920
RESULT 41			
US-09-495-406-6			
; Sequence 6, Application US/09495406			
; Patent No. 6503744			
; GENERAL INFORMATION:			
; APPLICANT: Gilbert, Michel			
; APPLICANT: Wakarchuk, Warren W.			
; APPLICANT: National Research Council of Canada			
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of			
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics			
; FILE REFERENCE: 019633-000110US			
; CURRENT APPLICATION NUMBER: US/09/495,406			
; CURRENT FILING DATE: 2000-01-31			
; PRIOR APPLICATION NUMBER: US 60/118,213			
; PRIOR FILING DATE: 1999-02-01			
; NUMBER OF SEQ ID NOS: 35			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 6			
; TYPE: DNA			
; ORGANISM: Campylobacter jejuni			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(876)			
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II			
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41			
US-09-495-406-6			
Alignment Scores:			
Pred. No.:	1.96e-167	Length:	876
Score:	1514.00	Matches:	279
Percent Similarity:	98.6%	Conservative:	8
Best Local Similarity:	95.9%	Mismatches:	4
Query Match:	96.3%	Indels:	0
DB:	3	Gaps:	0
US-10-734-719-9 (1-291) x US-09-495-406-6 (1-876)			
Qy	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGTTATTATTGCTGGAAATGGACCAGTTTAAAAAGAAATTTGATTATTCAAGA	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAATGATTTTGATGTTATTTAGATGCAATCAATTTTATTTTGAAGATAAACTACT	120
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	121	CTTGTAATAAAATGCAAGCAGTATTTTACAATCTAGTCTTTTGTGAACTACTAC	180

; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-09-816-028A-6

Alignment Scores:  
Pred. No.: 1.96e-167 Length: 876  
Score: 1514.00 Matches: 279  
Percent Similarity: 98.6% Conservative: 8  
Best Local Similarity: 95.9% Mismatches: 4  
Query Match: 96.9% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-09-816-028A-6 (1-876)

```
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCGTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTGATGATTTAGATGCAATCAATTTTATTGAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGACAGATTTTACAACTCTAGTCTTTTGTGAACATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGATATGAGACCGAATCAATCATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAACCAAGCTCATCTAGAAAATCAAAAATTTGTAAAAACTTTTACGATTTATTTTCCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTTCACAACTTAAAGATTCATGCTTATTATTA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCATCAATCAAGAAATTTACCTCAGGGTCTATATGTGCACAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTTGGGAAATTTGATTTTATCAAAAATGATCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACCAACAAAAAATCTTTTAAAAATGGCTCTCTAAATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACGAAATTAAGCTATATTGTTTATGCTCTTAACAGCTCTTTTAGCAAAATTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATCAAGAAAAAATACTACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATCTCATACCTTCTAGTGGCTTATGGAAAAATTTACAAAAATATTAATTTT 780
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATTAATAAAGAAATATTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
```

Db 841 CCTAGTGATATAAGCATTATTTCAAAGCAAAA 873

RESULT 43

US-10-303-162-6

; Sequence 6, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,162  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-10-303-162-6

Alignment Scores:

Pred. No.: 1.96e-167 Length: 876  
Score: 1514.00 Matches: 279  
Percent Similarity: 98.6% Conservative: 8  
Best Local Similarity: 95.9% Mismatches: 4  
Query Match: 96.9% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-162-6 (1-876)

```
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCGTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTGATGATTTAGATGCAATCAATTTTATTGAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGACAGATTTTACAACTCTAGTCTTTTGTGAACATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGATATGAGACCGAATCAATCATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAACCAAGCTCATCTAGAAAATCAAAAATTTGTAAAAACTTTTACGATTTATTTTCCCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATGATGATTTTTCACAACTTAAAGATTCATGCTTATTATTA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATTTCATCAATCAAGAAATTTACCTCAGGGTCTATATGTGCACAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
```

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGATCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAenLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAACAAATCTTTTAAATTTGGCTCCTTAATTTTAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATAATTCACATATATCGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 220  
Db 601 GAAAAAATTTACGAAATTAAGCTATATGTTATGTTCTTAACAGTCTTTTACGAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAenSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAATTTAAATTTTCAATTTTATCATCAAGAAATAAATACTACT 720  
QY 241 LysAspIleLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 260  
Db 721 AAAGATATACTCATACCTTCTAGTGGCTTATGGAAATTTTACAAAAATATTAATTTT 780  
QY 261 LysLysIleLysLysLysGluAenValTyrTyrLysLeuLysLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAATATTTTATCAAGTTGATAAAAGATCTTATTAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTTCAAAAGGAAAA 873

RESULT 44  
US-10-303-134-6  
; Sequence 6, Application US/10303134  
; Patent No. 6825019  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Michel  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,134  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR FILING DATE: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(876)  
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II  
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41  
US-10-303-134-6

Alignment Scores:  
Pred. No.: 1,966-167 Length: 876  
Score: 1514.00 Matches: 279  
Percent Similarity: 98.6% Conservative: 8  
Best Local Similarity: 95.9% Mismatches: 4  
Query Match: 96.9% Indels: 0  
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-134-6 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATCAAAAAAGTTTATTTATGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGA 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTGTATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120  
QY 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGGTAAAAAATGCAAGACAGATTTTACAACTCTAGTCTTTTGTGAAACAATACTAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATGAGACCGAAGCTAAATCACTGTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TTTTAAACCAAGCTCATCTAGAAAAATCAAAATTTTGTAAAACTTTTACGATTTATTTCCCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGAAATTCATGCTTATTTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACGAAATTTTATTTTCAATCAAGAAATTTACCTCAGGGGCTCTATATGTGCACAGTAGCC 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLysSerGlyIleAspPheTyrGluAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTTCTTTCCGGAAATTTGATTTTATCAAAATGGATCA 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAACAAATCTTTTAAATTTGGCTCCTTAATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATAAATTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTCAGCAATTAAGCTATATGTTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCAATTTTAAATTTTCAATTTTATCATCAAGAAATAAATACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATACTCATACCTTCTAGTGGCTTATGGAAATTTTACAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAATATTTTATCAAGTTGATAAAAGATCTTATTAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTATTTTCAAAAGGAAAA 873

RESULT 45  
US-10-303-118-6  
; Sequence 6, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21

```
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstii) from C. jejuni serotype O:41
US-10-303-118-6

Alignment Scores:
Pred. No.: 1.96e-167 Length: 876
Score: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservative: 8
Best Local Similarity: 95.9% Mismatches: 4
Query Match: 96.9% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-118-6 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGA 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTGATGATTAAGTCAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGACAGTATTTTACAATCCTAGTCTTTTTTGAACATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAATCAATGATGTCATAA 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAACCACGCTCATCTAGAAATCAAAATTTTGTAAAACCTTTTACGATTATTTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGAAATTCATGCTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnAlaValPheTyrThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACCAAAATTTATTTCATCAAAAGAAATTAACCTCAGGGGTCTATATGTGCACAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGATCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTGATACCAACCAAAAAATCTTTTAAAAATGGCTCTCTAATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACGAAATTAAGCTATATTTGTTATGTCCTAACAGTCCTTTTACGAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAAATACTATAC 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACATACCTTCTAGTGAGGCTATGGAATTTTACAAAAAATATTAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAATAATTAAGAAATATTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTTTCAAAGGAAAA 873

RESULT 46
US-10-303-128-6
; Sequence 6, Application US/10303128
; Patent No. 6911337
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstii) from C. jejuni serotype O:41
US-10-303-128-6

Alignment Scores:
Pred. No.: 1.96e-167 Length: 876
Score: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservative: 8
Best Local Similarity: 95.9% Mismatches: 4
Query Match: 96.9% Indels: 0
DB: 3 Gaps: 0

US-10-734-719-9 (1-291) x US-10-303-128-6 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGA 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTGATGATTAAGTCAATCAATTTTATTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATGCAAGACAGTATTTTACAATCCTAGTCTTTTTTGAACATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAATCAATGATGTCATAA 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAACCACGCTCATCTAGAAATCAAAATTTTGTAAAACCTTTTACGATTATTTTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTAAAGAAATTCATGCTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnAlaValPheTyrThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACCAAAATTTATTTCATCAAAAGAAATTAACCTCAGGGGTCTATATGTGCACAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGATCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTGATACCAACCAAAAAATCTTTTAAAAATGGCTCTCTAATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTTACGAAATTAAGCTATATTTGTTATGTCCTAACAGTCCTTTTACGAAATTTT 660
```

OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II ;  
OTHER INFORMATION: (cstII) from C. jejuni serotype O:4I  
US-10-735-419-6

Alignment Scores:	1.96e-167	Length:	876
Pred. No.:	1514.00	Matches:	279
Score:		Conservative:	8
Percent Similarity:	98.6%	Mismatches:	4
Best Local Similarity:	95.9%	Indels:	0
Query Match:	96.9%	Gaps:	0
DR:	5		

US-10-734-719-9 (1-291) x US-10-735-419-6 (1-876)

Qy	1	Meth	Lys	Lys	Vall	Ile	Ile	Ala	Gly	Asn	Gly	Pro	Ser	Ile	Lys	Glu	Ile	Asp	Tyr	Ser	Arg	20
Db	1	ATG	AAA	AAA	AGT	TAT	TAT	TGCT	GAAA	TGG	CAAC	AGT	TTT	AAA	AGG	AAA	TGAT	TAT	TTC	AAG	60	
Qy	21	Leu	Pro	Asn	Asp	Phe	Asp	Val	Phe	Arg	Cys	Asn	Gln	Phe	Tyr	Phe	Glu	Asp	Lys	Tyr	40	
Db	61	CTA	CAA	TAAT	GAT	TTT	TGAT	TAT	TAG	ATG	CAAT	CAA	TTT	TAT	TTT	TG	AAG	ATA	AA	TACT	120	
Qy	41	Leu	Gly	Lys	Cys	Lys	Ala	Val	Phe	Tyr	Thr	Pro	Asn	Phe	Phe	Phe	Glu	Gln	Tyr	Tyr	60	
Db	121	CTT	GCT	AAA	AAA	ATC	CA	AG	CAG	TAT	TTT	TACA	CA	TCT	AGT	CTT	TTT	TG	AA	CAAT	180	
Qy	61	Thr	Leu	Lys	His	Leu	Ile	Gln	Asn	Gln	Glu	Tyr	Glu	Thr	Glu	Leu	Ile	Met	Cys	Ser	80	
Db	181	ACT	TTT	AAAA	CA	TTT	TAAT	CC	AAAA	TCA	GAA	TAT	GAG	CCG	AA	TAAT	CAT	GTG	TCT	TAAT	240	
Qy	81	Tyr	Asn	Gln	Ala	His	Leu	Glu	Asn	Glu	Asn	Phe	Val	Lys	Thr	Phe	Tyr	Asp	Tyr	Phe	100	
Db	241	TTT	TA	ACC	AA	CG	AT	CA	TCT	AG	AAAA	TCA	AAA	TTT	TG	AAAA	CT	TTT	TAC	GAT	300	
Qy	101	Asp	Ala	His	Leu	Gly	Tyr	Asp	Phe	Phe	Lys	Gln	Leu	Lys	Glu	Phe	Asn	Ala	Tyr	Phe	120	
Db	301	GAT	GCT	CA	TTT	GG	AT	TAT	TTT	TCAA	CA	CA	CTT	AAA	GAAT	TCA	AT	TGC	TAT	TTT	360	
Qy	121	Phe	His	Glu	Ile	Tyr	Phe	Asn	Gln	Arg	Ile	Thr	Ser	Gly	Val	Tyr	Met	Cys	Ala	Val	140	
Db	361	TTT	CA	CG	AAA	TTT	TAT	TTT	CAAT	CA	AA	GAAT	TAC	CT	CAG	GGT	CT	TAT	ATG	CAC	420	
Qy	141	Ile	Ala	Leu	Gly	Tyr	Lys	Glu	Ile	Tyr	Leu	Ser	Gly	Ile	Asp	Phe	Tyr	Gln	Asn	Gly	160	
Db	421	AT	AG	CC	CT	TAG	GAT	CA	AA	GA	AA	TTT	AT	CTT	CG	GA	TTG	AT	TTT	TAT	480	
Qy	161	Ser	Tyr	Ala	Phe	Asp	Thr	Lys	Gln	Glu	Asn	Leu	Lys	Leu	Ala	Pro	Asp	Phe	Lys	Asn	180	
Db	481	TCT	TAT	GT	CTT	TG	AT	CA	CA	CA	AAAA	AA	CT	TTT	TAAA	AT	TGG	CT	CT	TAA	540	
Qy	181	Asp	Arg	Ser	His	Tyr	Ile	Gly	His	Ser	Lys	Asn	Thr	Asp	Ile	Lys	Ala	Leu	Glu	Phe	200	
Db	541	GAT	AAT	TC	AC	ACT	TAT	TC	GC	CA	TAG	TAAA	ATA	CAG	AT	TATA	AAA	AG	CTT	TGA	600	
Qy	201	Glu	Lys	Thr	Tyr	Lys	Ile	Lys	Leu	Tyr	Cys	Leu	Cys	Pro	Asn	Ser	Leu	Leu	Ala	Asn	220	
Db	601	GAAA	AA	CA	CTT	AG	AA	TAA	AG	CT	TAT	TG	TTT	TAT	GT	CT	CT	AA	CAG	CT	660	
Qy	221	Ile	Glu	Leu	Ala	Pro	Asn	Leu	Asn	Ser	Asn	Phe	Ile	Ile	Gln	Glu	Lys	Asn	Asn	Tyr	240	
Db	661	AT	AG	AA	CT	AG	CC	CA	AA	TTT	TAAA	TTT	TAT	CA	TAC	AA	GAAAA	AAAA	TAA	CT	720	
Qy	241	Lys	Asp	Ile	Leu	Ile	Pro	Ser	Ser	Glu	Ala	Tyr	Gly	Lys	Phe	Ser	Lys	Asn	Ile	Asn	260	
Db	721	AAA	GA	TAT	ACT	CA	TAC	TAC	CTT	CT	AG	GAG	CTT	TAT	G	AAAA	TTT	CA	AAAA	AA	780	
Qy	261	Lys	Lys	Ile	Lys	Ile	Lys	Glu	Asn	Val	Tyr	Tyr	Lys	Leu	Ile	Lys	Asp	Leu	Leu	Arg	280	
Db	781	AAAA	AA	AAAA	TAAAA	AA	TTT	AA	GA	AA	AT	TAT	TAT	TAC	AG	TTG	AT	AA</				

QY	81	Tyr	Asn	Gln	Ala	His	Leu	Glu	Asn	Gln	Asn	Phe	Val	Lys	Thr	Phe	Tyr	Asp	Tyr	Phe	Pro	100
DB	241	TTT	TAC	CAAG	CTC	AT	C	T	A	G	A	A	A	T	T	T	T	T	T	T	T	300
QY	101	Asp	Ala	His	Leu	Gly	Tyr	Asp	Phe	Phe	Lys	Gln	Leu	Lys	Glu	Phe	Asn	Ala	Tyr	Phe	Lys	120
DB	301	GA	TG	CTC	ATT	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	360
QY	121	Phe	His	Glu	Ile	Tyr	Phe	Asn	Gln	Arg	Ile	Thr	Ser	Gly	Val	Tyr	Met	Cys	Ala	Val	Ala	140
DB	361	TTT	TAC	GAAA	TTT	AT	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	420
QY	141	Ile	Ala	Leu	Gly	Tyr	Lys	Glu	Ile	Tyr	Leu	Ser	Gly	Ile	Asp	Phe	Tyr	Gln	Asn	Gly	Ser	160
DB	421	AT	AG	CC	CT	TAG	GAT	ACA	AAG	AAT	TAT	CT	T	T	T	T	T	T	T	T	T	480
QY	161	Ser	Tyr	Ala	Phe	Asp	Thr	Lys	Gln	Gln	Asn	Leu	Leu	Lys	Leu	Ala	Pro	Asp	Phe	Lys	Asn	180
DB	481	TC	T	T	AT	GC	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	540
QY	181	Asp	Arg	Ser	His	Tyr	Ile	Gly	His	Ser	Lys	Asn	Thr	Asp	Ile	Lys	Ala	Leu	Glu	Phe	Leu	200
DB	541	GATA	ATT	CAC	ACAT	TAT	CGG	CAT	AGT	AAAA	AAAT	CAG	ATAT	AAAG	CTT	TAG	AAAT	TTC	T			600
QY	201	Glu	Lys	Thr	Tyr	Lys	Ile	Lys	Leu	Tyr	Cys	Leu	Ser	Pro	Asn	Ser	Leu	Leu	Ala	Asn	Phe	220
DB	601	GA	AAAA	AACT	TAC	GAA	ATA	AAAG	CTAT	TAT	TG	T	T	T	T	T	T	T	T	T	T	660
QY	221	Ile	Glu	Leu	Ala	Pro	Asn	Leu	Asn	Ser	Asn	Phe	Ile	Ile	Gln	Glu	Lys	Asn	Asn	Tyr	Thr	240
DB	661	AT	AG	AACT	AG	CC	AAA	TTT	AAAA	TTT	T	T	T	T	T	T	T	T	T	T	T	720
QY	241	Lys	Asp	Ile	Leu	Ile	Pro	Ser	Ser	Glu	Ala	Tyr	Gly	Lys	Phe	Ser	Lys	Asn	Ile	Asn	Phe	260
DB	721	AA	GAT	AT	ACT	CT	T	AG	T	CAG	G	CT	TAT	TG	AAAA	TTT	T	T	T	T	T	780
QY	261	Lys	Lys	Ile	Lys	Ile	Lys	Glu	Asn	Val	Tyr	Tyr	Lys	Leu	Ile	Lys	Asp	Leu	Leu	Arg	Leu	280
DB	781	AAAA	AA	AAAA	AAAA	TTT	AAAA	AAAA	TTT	AT	T	T	T	T	T	T	T	T	T	T	T	840
QY	281	Pro	Ser	Asp	Ile	Lys	His	Tyr	Phe	Lys	Gly	Lys										291
DB	841	CCT	AGT	GT	AT	AA	AG	CAT	TAT	TT	CA	AG	CA	AAAA	873							

```

RESULT 47
US-10-735-419-6
; Sequence 6, Application US/10735419
; Patent No. 7026147
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Makarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Campylobacter Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)

```







Db 379035 AATCCTTACGCATTTTCATCATCAAAAGAAATATTATTAAATATTATACCTTCTTTTCA 378976  
Qy 180 AenAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199  
Db 378975 CAAATAAAGTCAAGCGCATATCCATCTATGGAATATGATTTAAATGCATTTATTT 378916  
Qy 200 LeuGluLysThrTyrLysIleGlyLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219  
Db 378915 TTACAAAACATTATGGAGTAATATTATTATGCAATTCGCCAGAAAGTCTCTATGTAAT 378856  
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsn 237  
Db 378855 TATTTTCCTTTATCACCA---CTGAATAACCAATTTACTTTTCTCGAAGAAAGAAA 378799  
Qy 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257  
Db 378798 AATTACACACAGATATTATTAATTCGCCGGAAGTTTGATATATATAAATAATGGT----- 378745  
Qy 258 IleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277  
Db 378744 ---ATATATCCAAACCAAGATTTACCAAAATCTGATTTTCGGTTGATCTGGGATATA 378688  
Qy 278 LeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 378687 TTACGTTTACCTAATGATATATAAAACACGCGCTTAAATCAAGA 378646

## RESULT 51

US-09-643-990A-1/c

; Sequence 1, Application US/09643990A

; Patent No. 6528289

; GENERAL INFORMATION:

; APPLICANT: Robert D. Fleischmann

; Mark D. Adams

; Owen White

; Hamilton O. Smith

; J. Craig Venter

; TITLE OF INVENTION: The Nucleotide sequence of

; the Haemophilus influenzae Rd Genome, Fragments

; Thereof, and Uses Thereof

; NUMBER OF SEQUENCES: 1

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville,

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3 1/2 inch diskette

; COMPUTER: Dell Pentium

; OPERATING SYSTEM: MS DOS v6.22

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/643,990A

; FILING DATE: 23-Aug-2000

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/487,429

; FILING DATE: 1995-06-07

; APPLICATION NUMBER: 08/426,787

; FILING DATE: 1995-04-21

; ATTORNEY/AGENT INFORMATION:

; NAME: Kenley K. Hoover

; REGISTRATION NUMBER: 40,302

; REFERENCE/DOCKET NUMBER: PBI86P1C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 301-610-5790

; TELEFAX: 310-309-8439

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1830121 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Alignment Scores:

Pred. No.: 4,55e-70 Length: 1830121

Score: 721.50 Matches: 145

Percent Similarity: 66.7% Conservative: 51

Best Local Similarity: 49.3% Mismatches: 89

Query Match: 46.2% Indels: 9

DB: 3 Gaps: 6

US-10-734-719-9 (1-291) x US-09-643-990A-1 (1-1830121)

Qy 2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21  
Db 379512 AAGTCTGTCATTTATGCGAGTAAATGGAACAAGTTTAAATCAATGACTATATATTATTA 379453  
Qy 22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41  
Db 379452 CCTAAGAGATTATGATGTTTCCGTTGCAATCAATTTATTTTGAAGATCATTATTTCTT 379393  
Qy 42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr 61  
Db 379392 GGCAGAGAAATATAAAAGTATTTTAAATTTGTTCTACAAATTTTGGACCAATCTATACT 379333  
Qy 62 LeuLysHisLeuIleGlnAsnGlnGluTyrGlu---ThrGluLeuIleMetCysSerAsn 80  
Db 379332 TTTATGCAATTAATTAATAAATTAATGAATATGAATATGCTGATATATTATTTATCATCTTTT 379273  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 379272 GTCAATTTAGGAGATTCAGATTAAGAAA---ATTAAAAATGTACAAAAATTTACTAACA 379216  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 379215 CAAGTTGATATGGACATTTATTTTAAACAAGCTACCGCCTTTGATGCTCTATTACAA 379156  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 379155 TATAACGAATATATGAAATTAAGAGAAATTTACATCAGGCGTTTATATGTGTGAGTGCA 379096  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGly 159  
Db 379095 ACTGTAATGGGTTATAAGATCTTTATTTTAAACAGGTATGATTTTATCAGAAAAAGGG 379036  
Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys 179  
Db 379035 AATCCTTACGCATTTTCATCATCAAAAGAAATATTATTAAATTTATACCTTCTTTTCA 378976  
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199  
Db 378975 CAAATAAAGTCAAGCGCATATCCATCTATGGAATATGATTTAAATGCATTTATTTT 378916  
Qy 200 LeuGluLysThrTyrLysIleLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219  
Db 378915 TTACAAAACATTATGGAGTAAATTTATTTGCAATTTCCCGAAGAAAGTCTCTATGTAAT 378856  
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsn 237  
Db 378855 TATTTTCCTTTATCACCA---CTGAATAACCAATTTACTTTTCTCGAAGAAAGAAA 378799  
Qy 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257  
Db 378798 AATTACACACAGATATTTTAAATTCGCCGGAAGTTTGATATATAAATAATGGT----- 378745  
Qy 258 IleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277  
Db 378744 ---ATATATCCAAACCAAGATTTTACCAAAATCTGATTTTTCGGTTGATCTGGGATATA 378688  
Qy 278 LeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 378687 TTACGTTTACCTAATGATATATAAAACACGCGCTTAAATCAAGA 378646

```
RESULT 52
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
```

```

; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature

```

```

Alignment Scores:
Pred. No.: 4,55e-70 Length: 1830121
Score: 721.50 Matches: 145
Percent Similarity: 66.7% Conservative: 51
Best Local Similarity: 49.3% Mismatches: 89
Query Match: 46.2% Indels: 9
DB: 3 Gaps: 6

```

US-10-734-719-9 (1-291) x US-10-158-865-1 (1-1830121)

```

Qy 2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
Db 379512 AAGTGTGATTTTCAGGTATGGAACAAGTTAAATCAATTGACTATAGTTTATTA 379453
Qy 22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41

```

```

Db 379452 CCTAAAGATTATGATGTTTCGTTGCAATCAATTTATTTTGAAGATCATTATTTCTT 379393
Qy 42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr 61
Db 379392 GGCAAGAAAAATAAAAAAGTATTTTAAATGTTCTACAAATTTTTCAGCAATACTACT 379333
Qy 62 LeuLysHisLeuIleGlnAsnGlnGluTyrGlu---ThrGluLeuIleMetCysSerAsn 80
Db 379332 TTTATGCAATTAATTAATAATAATGAATATGAATATGCTGATATTTATTTATCATCTTT 379273
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 379272 GTCAATTTAGGAGATTTCAGAAATTAAAGAAA---ATTAAAAATGTCACAAAAATTAACA 379216
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 379215 CAAGTTGATATTGGACATTATTATTAAACAAGCTACCCGCCCTTTGATGCTATTTACAA 379156
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 379155 TATAACGAATTATATGAANAATAAGAAATTATCATCAGCGCTTTATATGTCAGATGGCA 379096
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGly 159
Db 379095 ACTGTAATGGGTATAAAGATCTTTATTAAACAGGTATTGATTTTATCAAGAAAAAGGG 379036
Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProaspPheLys 179
Db 379035 AATCCTTACGCATTTTCATCAAAAAGAAAAATATTATAAATTTATACCTCTTTTCA 378976
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 378975 CAAAATAAAGTCAAAGCGATATCCATCTCTATGGAATATGATTTAAATGCACCTTTATTT 378916
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 378915 TTACAAAAACATTATGGAGTAAATATTATTCATTCGCCAGAAAGTCTCTATGTAAT 378856
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsn 237
Db 378855 TATTTTCCTTTTATCACCA--CTGAATAACCCCAATTACTTTTATTCTCGAAGAAAAAGAA 378799
Qy 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257
Db 378798 AATTACACACAGATATTTTTAAATTCGCCGGAAGTTTGTATATAAAAAAATTCGT----- 378745
Qy 258 IleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277
Db 378744 ---ATATATCCAAACCAGAAATTTACCAAAATCTGATTTTTCGGTTGATCTCGGATATA 378688
Qy 278 LeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
Db 378687 TTACGTTTACCTAATGATATATAAACACGCCCTTAAATCAAGA 378646

```

# RESULT 53

```

US-09-113-750A-33
; Sequence 33, Application US/09113750A
; Patent No. 6294176
; GENERAL INFORMATION:
; APPLICANT: David E. Junker and Mark D. Cochran
; TITLE OF INVENTION: Recombinant Raccoonpox virus
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/113,750A  
 FILING DATE:

CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.  
 REGISTRATION NUMBER: 28,678  
 REFERENCE/DOCKET NUMBER: 55744  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)262-0400  
 TELEFAX: (212)664-0525  
 TELEX: 422523

INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2375 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..738  
 FEATURE: CDS

NAME/KEY: CDS  
 LOCATION: 799..2375  
 FEATURE: CDS

NAME/KEY: unsure  
 LOCATION: 1106  
 FEATURE: unsure

NAME/KEY: unsure  
 LOCATION: 1109  
 FEATURE: unsure

US-09-113-750A-33

Alignment Scores:  
 Pred. No.: 0.00409 2375  
 Score: 113.50  
 Percent Similarity: 37.0%  
 Best Local Similarity: 20.1%  
 Query Match: 7.3%  
 DB: 3

US-10-734-719-9 (1-291) x US-09-113-750A-33 (1-2375)

Qy 24 AspPheAspValPheArgCys-----AsnGlnPhe 33  
 |||||  
 Db 577 GATTTCGATGATATCGGTTTATATAGTGAAGGTGGTGTATCACCAGGTCGACATTG 636  
 Qy 34 TyrPheGluAsp-----LysTyrTyrLeuGlyLys-----LysCysLysAlaValPhe 49  
 |||||  
 Db 637 TATTTCGAAATGTTACTGAAGAATATAGGTAATACATATACCTGTGCGGGA----- 690  
 Qy 50 TyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnAsnGln 69  
 |||||  
 Db 691 -----CACAACTATTATTTTCAGAAACCCCTTACGACCACAGTAGTATTG----- 735  
 Qy 70 GluTyrGluThrGluLeuLeuMetCysSerAsnTyrAsn-----GlnAla 84  
 |||||  
 Db 736 GAGTAAACAACTAACATTTTATATCTGAAATTTAACGTATATCTCCTGTTCTATAT 795  
 Qy 85 HisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAsp---AlaHis 103  
 |||||  
 Db 796 AAAATGGATGAAGATAGGATGCTACTAAGTATATGTTATCTCACTGATAGAGAACAT 855  
 Qy 104 LeuGlyTyrAsnPhePheLysGlnLeuLysGluPhe-----AsnAlaTyrPheLys 120  
 |||||  
 Db 856 ATAAATATAGACTCTGTGTTAAACGGTTATGTGAATAATCGGATCTTAATGCGTGTATAGA 915  
 Qy 121 -----PheHisGluLeuTyrPheAsnGlnArgGileThrSerGlyValTyr 135  
 |||||

Db 916 TCAGGATGTACCGCTTTTACACGAGTACTTTTATAAATGTAGATCAATCAAAAGGAAAAATAC 975  
 Qy 136 MetCysAlaValAlaLeuLeuGlyTyrLysGluLeuTyrLeuSerGly---IleAsp 154  
 |||||  
 Db 976 GATTATAGGTATAAT-----GGTTATTATAAATATTATTATTCAGGAGATTATGAA 1026  
 Qy 155 PheTyrGlnAsnGlySerSerTyrAlaPheAsp----- 165  
 |||||  
 Db 1027 AACTATAACAGTAATTATGATGATGCTATAGATATGACACGATGACGACCGGAAGAT 1086  
 Qy 166 -----ThrLysGlnGluAsn----- 170  
 |||||  
 Db 1087 GATGAATCATCTCGAAACAANATNTGAATTTTATGATGAACACACAGATCAAAATATCAG 1146  
 Qy 171 -----LeuLeuLysLeuAlaProAspPheLysAsnAspArgSer 183  
 |||||  
 Db 1147 CTAATATGTTCTAATATTAACCTCACAAATAGAGTAGATGATTTCTGTAGATGAA----- 1200  
 Qy 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr 203  
 |||||  
 Db 1201 ---TTCATGCTTACGATCAGATATCGACATAAAT-----GATTATGTAGATAAATCT 1251  
 Qy 204 TyrLysIleLysLeuTyr-----Cys-----Leu 211  
 |||||  
 Db 1252 ATAAACATCTAGTATATGGAAGCGATAATAGGCGCAATCGATGGTGTGATGTATGGAGA 1311  
 Qy 212 CysProAsnSer----- 215  
 |||||  
 Db 1312 TGTCACCATAGCGGATCATACAGATTCGGTAAGGAATGTATAGACAATATATATGAAGAT 1371  
 Qy 216 -----LeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn 230  
 |||||  
 Db 1372 AATAATACAAAAATAGAAATTCGATAATTCGATTCGTTGTCACAACTTACTGTATCGCGAA 1431  
 Qy 231 PheIleIleGlnGluLysAsnAsnTyrThr-----LysAspIleLeu 244  
 |||||  
 Db 1432 TATGTAATTCGGTAACTAGAGATCGCTACACAAATATGGGAAAAGAAATCAGTGTTA 1491  
 Qy 245 IleProSerSerGluAlaTyr-----GlyLysPheSerLysAsnIleAsnPheLysLys 262  
 |||||  
 Db 1492 GATAGATACATGAGTGCATATAGCGGTAAACCGATATAGTAAACATTTCTCTTTAAGGGA 1551  
 Qy 263 IleLys-----IleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeu 278  
 |||||  
 Db 1552 TTTTCTGATTACGTTAGAAAAAATGATTTAGATATGATTTGTTTAAAGAACTACTTT 1608

## RESULT 54

US-09-662-254B-27  
 ; Sequence 27, Application US/09662254B  
 ; Patent No. 6933145  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moyer, Richard W.  
 ; APPLICANT: Li, Yi  
 ; APPLICANT: Bauden, Alison Louise  
 ; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
 ; FILE OF INVENTION: Vertebate Cells  
 ; FILE REFERENCE: UF-221C1XC1  
 ; CURRENT APPLICATION NUMBER: US/09/662,254B  
 ; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 09/086,651  
 ; PRIOR FILING DATE: 1998-05-29  
 ; PRIOR APPLICATION NUMBER: 60/224,479  
 ; PRIOR FILING DATE: 2000-08-10  
 ; NUMBER OF SEQ ID NOS: 80  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 27  
 ; LENGTH: 32392  
 ; TYPE: DNA  
 ; ORGANISM: Amsacta moorei entomopoxvirus  
 ; US-09-662-254B-27

Alignment Scores: 0.194 Length: 32392  
 Pred. No.:

```
Score: 113.00 Matches: 79
Percent Similarity: 33.3% Conservative: 56
Best Local Similarity: 19.5% Mismatches: 100
Query Match: 7.2% Indels: 170
DB: 3 Gaps: 20

US-10-734-719-9 (1-291) x US-09-662-254B-27 (1-32392)

QY 13 LeuLysGluLeuAspTyrSerArgLeu-----ProAsnAspPheAspVal 27
DB 12394 ATGAAGAATAATTATATTCTCAACTATCATATACATATACACCCCGATATTTATATAA 12453

QY 28 PheArgCysAsnGlnPheTyrPheGlu-----AspLysTyrTyrLeuGlyLysLys--- 44
DB 12454 TTTAGA-----AATTATTATAGAAATTAATACGAATATATATCCCTAAATAACT 12507

QY 45 -----CysLys 46
DB 12508 TTGATAATTAATAAAAAATAAGTATTACTATAATAATAATAATAATTTAGACTTAAAA 12567

QY 47 AlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 66
DB 12568 ACTATCGAATATATCAATAACGAATCAGAGACATATATGATAACATAATACACATTCAT 12627

QY 67 GlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeu 86
DB 12628 AGAAAT-----TTATCTGTTAATTTAATTAATAATAATAATATATATATATATAT 12666

QY 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAsnAlaHisLeuGlyTyr 106
DB 12667 -----TAT 12669

QY 107 AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys-----PheHisGlu 123
DB 12670 TATATTTTGATACAGAAAATATATATATATATATATATATATATATATATATATATATAT 12729

QY 124 IleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal----- 139
DB 12730 ATATATTACGATACATCTATTTCTTATTAATTAATTAATTTTGCATATATACACACACGAT 12789

QY 139 ----- 139
DB 12790 AATAAATTAATAATATTATAGTCGTATATATACACATGCATTAAATGACAAGTATTCAT 12849

QY 140 -----AlaIleAlaLeuGlyTyrLysGlu--- 147
DB 12850 AAAAAATTAATATCTCGATTGTTTATTTGATGTTTATCTTTTAAATAATAAAAAATGT 12909

QY 148 -----IleTyrLeuSerGlyLysPhe----- 155
DB 12910 AACAAAGATTCCTATTATTATTTTAAACCAAAATTTTACAATCTAGATACAAATAAT 12969

QY 156 -----TyrGlnAsnGlySerSerTyrAlaPheAsp 165
DB 12970 TCAATCGTAATATTGATTGATAATTCGTATTATATAGAGGAAATGCA---CTTATTGAG 13026

QY 166 ThrLysGlnGluAsnLeuLeuLysLeuAlaProAsp----- 177
DB 13027 TTCCTAGATAAAAAATAATTTAAAAATAATTAATGATAAATTTTATCTAATAATAACTAAT 13086

QY 178 -----PheLysAsnAspArgSerHisTyrIleGly 187
DB 13087 AATTGGATTGATATATTATGGAAACAAAAATTTAAAAATTCATTGAATAATATATCAAT 13146

QY 188 HisSerLysAsn-----ThrAsp 193
DB 13147 TATTGTAGTATTATTATATAATAAATGATAATTTATTTATACACAAATGAATAACAGAT 13206

QY 194 IleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyr---CysLeuCys 212
DB 13207 -----AAATATATAGATAGAAATTAATAATAATAATAATAATAATTCGATAATGTGTGT 13254

QY 213 ProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIle 232
DB 13255 AAAGGTAATAATAAT-----ATAGAACAATATATGATCGCAATCTACATTTATA 13305

QY 233 IleGlnGluLysAsnAsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTyrGly 252
DB 13306 CTTAATAAAGATAATATATATATATATATATATATATATATATATATATATATATAT 13350

QY 253 Lys-----PheSerLysAsnIleAsnPheLysLysIleLysIle 265
DB 13351 AAATATATGTTTAAATAACGAATCTTTTATTAAAAATATTAAT---AGAAAAAGACCAAT 13407

QY 266 LysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLys 285
DB 13408 ATTCAATAATTATCAATTATCAATGTTAGATAATAATAGTTTA---TTAAATCATTTAAAT 13464

QY 286 HisTyrPheLysGly 290
DB 13465 AATATTTTAAAGGT 13479

RESULT 55
US-09-790-988-1/c
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRO
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Alignment Scores:
Pred. No.: 13.7 Length: 640681
Score: 113.00 Matches: 74
Percent Similarity: 35.5% Conservative: 57
Best Local Similarity: 20.1% Mismatches: 112
Query Match: 7.2% Indels: 126
DB: 3 Gaps: 20

US-10-734-719-9 (1-291) x US-09-790-988-1 (1-640681)

QY 3 LysValIleIleAlaGlyAsn-----GlyProSerLeuLysGluIleAspTyrSerArg 20
DB 240894 AAAGCTTTTCTAGGTGGTAAATATAGGTGTTCCGTACTAGAAATACCTGATAAGAAAGCA 240835

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 240834 -----GATTATACATAATAAGAACTGCTAGTTTTCACATAGAAAAATACATTTAAT 240784

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThr-----Pro 52
DB 240783 TTTAAATCTAAGATAGACACTTATTTCTTAATATTAGCGAAGATCATATCAATCGATACCCA 240724

QY 53 AsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72
DB 240723 AAT---GGATTCCAACATAATAAAACACATAAATTCCTCTCTTTTATAATCAA----- 240676

QY 73 ThrGluLeuIleMetCysSerAsnTyrAsnGln-----AlaHisLeuGlu 87
DB 240675 GCAGAAAATTTGATAATTAATACTTAATGATAAGATAAGAAAAAGTCTTTATTCATAGTAAG 240616

QY 88 AsnGluAsnPheValLysThrPheTyrAspTyrPheProAsnAlaHisLeuGlyTyrAsp 107
```

```
Db 240615 AATAAAATGGATAGT-----TTTGAACAATAGAGTATATCGT 240571
Qy 108 PhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn 127
Db 240570 ATTGTGTTCTAAAGCAATGACCTATTTATTTTAA-----AAT 240529
Qy 128 GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGlu 147
Db 240528 AAAAAATA-----TTAAATACATAGTCAA 240505
Qy 148 IleTyrLeuSerGlyIleAspPheTyrGlnAsn-----GlySerSerTyr 162
Db 240504 ATATTATATGATGATATCATATAATTAATATATATAGTCTCTTTAGCTATTTTCAGAT 240445
Qy 163 AlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPhe----- 178
Db 240444 GCAATGCAATTCCTAGAAATGACGAATAATGTAATCTTAAAGTTTCTCGAATTTACCA 240385
Qy 179 -----LysAsnAspArgSer---HisTyrIleGlyHisSerLysAsn 191
Db 240384 CATCGATTTCAAATPAATAAAAAATGAAAAAGGAGTACGTTGGATAAACGATTTCTAAATCC 240325
Qy 192 ThrAspIleLysAla----- 196
Db 240324 ACAATGTAAATAGTACTCAAGTTGCTTTAAATTTCTATATAAAACAACCTGGAAACAATAAGA 240265
Qy 197 -----LeuGluPheLeuGluLys 202
Db 240264 TTATTATTAGCGCGGATAGTAATCTGCAAAATTTTAATATATTAAAAAATATTTTAGA 240205
Qy 203 ThrTyrLysIleLysLeuTyrCys-----LeuCys 212
Db 240204 ACTTTAAAAATAAAAAATTTACTGTTTGGAAAGAGATGCAATTAATTTGTCAAAAATATGT 240145
Qy 213 ProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIle 232
Db 240144 GAAAAAAATCTATT-----TATGTGAA-----AATTTAAAAAAGCAGTAAAT 240100
Qy 233 IleGlnGluLysAsnAsnTyrThrLysAsp-----IleLeuIleProSerSerGluAla 250
Db 240099 TTAATTTCAAACAACAGTCAATACAGGTATACCGTTCTTTTGTCTCTCGGTGCGAGCAGC 240040
Qy 251 TyrGlyLysPheSerLysAsnIleAsnPheLysIleLysIleLysGluAsnValTyr 270
Db 240039 TTAGGTCAATTTTCT-----AACCTTGAAGAG-----AGAGGTAATCTTTT 239998
Qy 271 TyrLysLeuIleLysAspLeuLeu----- 278
Db 239997 ATAAAAATTAATAAGGAATAACTTGAACACACTTACAAAAAATAATTCAAAAATCCAA 239938
Qy 279 ---ArgLeuProSerAspIleLysHis 286
Db 239937 AAAAGATTGAAGAAAAATATACACAC 239911
```

## RESULT 56

```
US-08-545-528D-1/c
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PBI93P1
; CURRENT APPLICATION NUMBER: US/08/545,528D
; PRIOR FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 1
; LENGTH: 580073
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1
Alignment Scores:
Pred. No.: 13.6 Length: 580073
Score: 112.50 Matches: 61
Percent Similarity: 36.7% Conservative: 40
Best Local Similarity: 22.2% Mismatches: 87
Query Match: 7.2% Indels: 87
Gaps: 13
DB: 3
```

US-10-734-719-9 (1-291) x US-08-545-528D-1 (1-580073)

```
Qy 60 TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSer 79
Db 440508 TATTTCAATAAATACCTAATTAAGAACCAATAATTTTGCAATCATCTGTTATCTTAAACAGT 440449
Qy 80 AsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAsp----- 97
Db 440448 GAACACCAAAATGTTTAATAGTAATAATATGAAGCTTTTGTAAAGCAATGGATGATATCCTG 440389
Qy 98 -----TyrPheProAspAlaHis----- 103
Db 440388 AATCATATAAATAAGTAAAAATTTTAAACAATGCTCTCATAAAAATTAATGATCAAAACC 440329
Qy 104 -----LeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsn 116
Db 440328 CTCGGTAATGAGCAACGCTATTAGCTTATGATTTTTCAAAAATAGCAATAAGGTTGGA 440269
Qy 117 -----AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThr 131
Db 440268 CTTTATTCAACATATACATATCTAGATTTTGTAAATTTTATTAGTTAGATCTTAAAGTA 440209
Qy 132 SerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSer 151
Db 440208 AATTATCTGTTGTTTAAACAAGTAGTCTTCCAATA---TATCAAAAGATTTATTATTGAT 440152
Qy 152 GlyIleAspPhe-----TyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGlu 169
Db 440151 AACCTCCCTTTTATTAGTTTTCGATAATAACTTGTGAAGTGCAATTTGGT----- 440104
Qy 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHis---TyrIleGlyHis 188
Db 440103 -----TTAGCATTAAGAAAAATAAAAGTCAAAACGATACTGTATTTCGTTTGTGAAAAA 440050
Qy 189 SerLysAsnThrAspIleLysAlaLeuGluPheLeuGlu-----LysThrTyrLys--- 205
Db 440049 ACCAAAAATACAGATACAGAAATTAATAGTTTAAAAAATACTGAAAACTTTTAAGGGT 439990
Qy 206 IleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaPro 225
Db 439989 TTGAAAAATAGTTTTTTTATTAATTCA----- 439960
Qy 226 AsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIle 245
Db 439960 ----- 439960
Qy 246 ProSerSerGluAlaTyrGlyLysPheSerLysAsnIle-----AsnPheLysLysIle 263
Db 439959 CCAGAAGAAAAACAACACTATTAAAGCTTTCGTATTCGTAATTAAGAAATTAATAAACAA 439900
Qy 264 LysIleLysGluAsnValTyr-----TyrLysLeuIle 274
Db 439899 AAAATAAAACATGAAGTTTATTTCTTTAAGATCATTTCAAAATAAGTTTTTTTAAGTTAGTT 439840
Qy 275 LysAspLeuLeuArgLeuProSerAspIleLysHisTyrPheLys 289
Db 439839 AGAAAAATTA-----GAAAAAAGCACAGTCTATAA 439810
```

RESULT 57



US-09-492-709A-188  
; Sequence 188, Application US/09492709A  
; Patent No. 6720139  
; GENERAL INFORMATION:  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Trawick, John  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Froelich, Jamie M.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
; FILE REFERENCE: ELITRA.001A  
; CURRENT APPLICATION NUMBER: US/09/492,709A  
; CURRENT FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 485  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 188  
; LENGTH: 1179  
; TYPE: DNA  
; ORGANISM: E. Coli  
US-09-492-709A-188

Alignment Scores:  
Pred. No.: 0.00226 Length: 1179  
Score: 112.00 Matches: 70  
Percent Similarity: 36.1% Conservative: 39  
Best Local Similarity: 23.2% Mismatches: 85  
Query Match: 7.2% Indels: 108  
Gaps: 17

US-10-734-719-9 (1-291) x US-09-492-709A-188 (1-1179)

Qy 25 PheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu----- 41  
Db 259 TTGACTCTATCGAATAAATAGATTGGTAGATTGGAAGATATTATATATACACTA 318

Qy 42 ---GlyTyrLysCysLysAlaValPheTyr----- 50  
Db 319 AGAGCGCGCTTTTATCGCTATTATCATCACCAGATCAATTGAAACCAACTGGATACAGAA 378

Qy 51 ThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnGlu 70  
Db 379 ACAATAACTATTAGTTATGAA---TTTATCTCTGTCAACATCTTATCGCAAAATACCGAG 435

Qy 71 TyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsn----- 88  
Db 436 TTC---ACGTATTGGCACTCCCTGAAATATTAAACCGTTGTTTTCACCAATTCAAAA 492

Qy 88 ----- 88  
Db 493 AACCAACAAATAATCGTTTCAAAACGTTAAACAGCAAGCTATTGGCAGGCTACTTGTCT 552

Qy 89 -----GluAsnPheValLysThrPheTyrAspTyrPhe 99  
Db 553 GCTGTGCGGTATACAATGGGAACATTGAGGATTCAGAGATCTCTCGCGGAAAAAATCGGT 612

Qy 100 ProAspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPhe 119  
Db 613 GGAGATGCAATAAAGGCTATGAT-----CAATA----- 642

Qy 120 LysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal 139  
Db 643 -----CTAAATGAAAAA---ACAGCGGCATAGCATAGCAACACACA 681

Qy 140 AlaIleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159  
Db 682 TCTATTCTTTTAAACAAGCGT-----TCTAATGTTGATACATATACAGAAATA 729

Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLys 179  
Db -----

Db 730 AATAGTTAC-----TTAGGCAAACTTAGA----- 753  
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199  
Db 754 -----GGTCAACAAAACTTCTTGAT-----GGTATAGACATA 786  
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219  
Db 787 ATAGAA-----ATAATATACATTAGAGACCTTCAAAAGACTTAGCTAAC 831  
Qy 220 Phe---IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsn 238  
Db 832 TTACGAAAGGAGCTTTAATAAAAACTGTAGAAAAAATTTCTTATCAAA----- 879  
Qy 239 TyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn--- 257  
Db 880 -----CTTGCAAAAACCTCCGAAGCATCTGGAGATTCACGCCGGAAGAC 924  
Qy 258 ---IleAsnPheLysLys-----IleLysIleLysGluAsnValTyrTyrLysLeuIle 274  
Db 925 CTTTAAAGAATCGAAGAGGCAATGTTCTCTCTAAATTAATAATGTTCCACCAATAACTATCT 984  
Qy 275 LysAsp 276  
Db 985 CTAGAT 990

RESULT 58  
US-08-852-629-11  
; Sequence 11, Application US/08852629  
; Patent No. 6106825  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W  
; APPLICANT: Li, Yi  
; APPLICANT: Hall, Richard L  
; TITLE OF INVENTION: ENTOMOPOXVIRUS-VERTEBRATE GENE DELIVERY  
; TITLE OF INVENTION: VECTOR AND METHOD  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: Florida  
; COUNTRY: U.S.A.  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/852,629  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bencan, Gerard H  
; REGISTRATION NUMBER: 35,746  
; REFERENCE/DOCKET NUMBER: UF-184  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 407-426-7500  
; TELEFAX: 407-839-8589  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4810 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-852-629-11

Alignment Scores: 0.0168 Length: 4810  
Pred. No.:



```
QY 99 PheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyr 118
Db 4293 -----TTAGAAGAA----- 4301

QY 119 PheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAla 138
Db 4302 -----TTGCATATAATATATATGATATAATAATATTTTAAATAAT----- 4340

QY 139 ValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsn 158
Db 4341 -----ATTCCAGAAAATATTAAAGTTTATATATTTCAAAATTTA-----AAT 4382

QY 159 GlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPhe 178
Db 4383 ATTATTAATTTAAATTTTATAACAAAATTAATAATAATAACATATTTAGAT----- 4433

QY 179 LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLys----- 195
Db 4434 -----ATATCTTATAACAAAATAGCAATATAAGTAATATTATA 4472

QY 196 -----AlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCys 212
Db 4473 CTACCACATCTCTAGAAATTTTAAAT-----TGTCGAATCA 4508

QY 213 ProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIle 232
Db 4509 TGTAAATAAATAGCATATAATTTTATTAATAATTTAGTAAATTTAAATAATAATA 4568

QY 233 IleGlnGluLys-----AsnAsnTyrThrLysAspIleLeuIleProSer 247
Db 4569 TCTAAATAAATTTGGTAACCTTTAATAATGTTTTCTTATAGTATAGTGAATTAAT 4628

QY 248 SerGluAla-----TyrGlyLysPheSerLysAsnIleAsnPheLysLys 262
Db 4629 ATGGAATCAATACAAATAAAAGATTAATAATTTATAGAAAATTAATTAATTTAAATA 4688

QY 263 IleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSer 282
Db 4689 TTAGATATATCTTCAATGTTTAAATAAAATATATA---CATTTGATAAAATTTCCAAA 4745

QY 283 AspIleLysHis 286
Db 4746 AGTATAACTCAT 4757

RESULT 60
; US-09-662-254B-25/c
; Sequence 25, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bowden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; TITLE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UF-221C1XC1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
; US-09-662-254B-25

Alignment Scores: 0.471 Length: 50000
Pred. No.: 112.00 Matches: 72
Score:
```

Percent Similarity: 40.6% Conservative: 41  
Best Local Similarity: 25.9% Mismatches: 123  
Query Match: 7.2% Indels: 42  
DB: 3 Gaps: 15

US-10-734-719-9 (1-291) x US-09-662-254B-25 (1-50000)

QY 34 TyrPheGluAspLysTyr-----TyrLeuGlyLysLysCysLysAlaVal 48  
Db 39859 TATTTTACGAGAGATATATATTTTATATACATATTTG-----ACA 39818

QY 49 PheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsn 68  
Db 39817 TTTTATTTCT---AATATTATATACATATCAAGATCGAGAAAATAATATTTTAAAAAAT 39761

QY 69 GlnGluTyrGluThrGluLeuIleMet-----CysSerAsnTyrAsnGlnAla 84  
Db 39760 AGAGTATTGCAATTTACCAGAGTTATAAATAATATATATATATATATATATAAATTTA 39701

QY 85 HisLeuGluAsn-----GluAsnPheValLysThrPheTyrAspTyrPheProAsp 101  
Db 39700 AATATTGAAAATATTGTTATAGATCTATTAGAAAANNTTCAATATAATAATAATGTA 39641

QY 102 AlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys-- 120  
Db 39640 CAAGATTTGATATATAGT---TTAGATCAGATCAAGAATTTTATAATAATAATAAAT 39584

QY 121 ---PheHisGluIleTyrPheAsnGlnArg-----IleThrSerGlyValTyrMetCys 137  
Db 39583 TATATACATGATATTATTTAAAAATGCAATAATATTATGATCAATTAATTTCCAGAATA 39524

QY 138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln 157  
Db 39523 CCATTTGATATTATATAAAAAATAAGAAATTTATTGATAAATAACTAATATTATATTAAA 39464

QY 158 AsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAsp 177  
Db 39463 AAT-----TATATATCAGATGATAAATTAGTAAATATTATATCTAAAAATCCATTA 39413

QY 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu 197  
Db 39412 TTTTAAATAATATTAAGAATATATA-----AATGAAAATAATAAAATTAAT 39365

QY 198 GluPheLeuGluLys-----ThrTyrLysIleLysLeuTyrCysLeuCysProAsnSer 215  
Db 39364 AATATTATTAATAAAATAGTAATACTAATGATGATTTTATATTTATTTATATAATATAAT 39305

QY 216 LeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlu 235  
Db 39304 AAAAATTCGATTTTATATAAATTTATTTAAAAAT---AATGAAGATTTTATTTTATTTT 39248

QY 236 LysAsnAsnTyrThrLysAspIleLeu-----IleProSerSerGluAlaTyrGlyLys 253  
Db 39247 AAGAATAATAAATAAATGATTTATATATATATAATATATATATATCATCGGATGAATAT 39194

QY 254 PheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeu 273  
Db 39193 -----AAAAATATAATTTGATATGGAATAAATTTAAATCAAAAAATATA-----GATATT 39146

QY 274 IleLysAspLeuLeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 39145 ATAATGGATTTATGTAAGAAAAAATAGCGATGTTTATATAAAAAATTAACAAAAA 39092

RESULT 61  
US-09-601-198-153  
; Sequence 153, Application US/09601198  
; Patent No. 6531583  
; GENERAL INFORMATION:  
; APPLICANT: Cassell, Gail H.  
; APPLICANT: Chen, Ellison Y.  
; APPLICANT: Glass, Jennifer S.  
; APPLICANT: Glass, John I.  
; APPLICANT: Helner, Cheryl R.

```
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 153
; LENGTH: 1866
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-153

Alignment Scores:
Pred. No.: 0.00497 Length: 1866
Score: 111.50 Matches: 67
Percent Similarity: 36.5% Conservative: 43
Best Local Similarity: 22.3% Mismatches: 114
Query Match: 7.1% Indels: 77
DB: 3 Gaps: 14

US-10-734-719-9 (1-291) x US-09-601-198-153 (1-1866)
Qy 16 IleAspTyrSerArgLeuProAsn-----AspPheAspValPheArgCysAsnGln 32
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
736 ATTGATTATCATCGAAGAGATAATTTTAAAAAGATATTTTAGATTATAAAAAAAT--- 792
Qy 33 PheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyr----- 50
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
793 -----TATGAAGATTATTATATATATATATATGATCAAAAGCTTAACCAAGATTATCAACAG 846
Qy 51 -----ThrProAsnPhePheGluGlnTyr 59
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
847 GGATTAGATAAATAAATAAGATAATATATGAAGTCAATATTTTATTTT-----TTT 900
Qy 60 TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSer 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 AATACTAAAAAACCCACTTTTCAAAAATTTAGTTATCAAAATGTTTTTAGAATTAAGAAT 960
Qy 79 ----- 79
Db ----- 79
Qy 961 AAGTTGGGTTTAAACGTTATGTTTTTAAGAAATATTCATCAAAAAGAGTTAGAGCACGT 1020
Qy 80 -----AsnTyrAsnGlnAlaHisLeuGluAsnGlnAsnPheValLysThrPheTyr 96
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1021 GATCATTTTATGCAATGAAAAAACAACAAATAGAAATTTATTTAGATCTTTT--- 1077
Qy 97 AspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsn 116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1078 -----GATAAAGAAATTCGAATAT-----TTTCAAAAATATATTTGATCATATAAT 1119
Qy 117 AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMet 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1120 TTAGATATTAATATAATAAATTTCAAAAGGAGCATATATTACACAAAAAATAATAAA 1179
Qy 137 CysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr 156
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1180 AATCTTTTATGATGCTTTCATAATAATGATGTTCAAAATTACAAAATATCTTAATTCAA 1239
Qy 157 GlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaPro 176
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1240 CAAAATGATAGGAATAC-----CAAAATAAATTAATTAACAAATTAATA 1284
Qy 177 AspPheLysAsnAspArgSerHisTyrIleGlyHisSerLys-----Asn 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1285 GAATTAGTAAT-----TATTTTAAAAACCAAAAAACACCAAGAGATTCTAAT 1332
Qy 192 ThrAspLeuAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1333 GAATATAAGCAAAATTAATGATTATGAAAAAATATTTTTCACAAAATTTGGTAT----- 1386
```

```
Qy 212 CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1387 -----TTTATTTTAAGAAATTTAAAGACACTTGAATTAAGTTTAAATAATTGATT 1437
Qy 232 IleIleGlnGluLys---AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAla 250
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1438 ACITTACCAAAACAAATAAATCCTATTACTAAAGAGATGTTAGGACCAAGTGAA----- 1491
Qy 251 TyrGlyLysPheSerLysAsnIleAsnPhelysLysIleLysIleLysGluAsnValTyr 270
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1492 -----AAATATTTAAAAAGGTTA-----AAAAAACTTAAACCATATATGAGGATTATTAT 1539
Qy 271 Tyr 271
Db 1540 TTC 1542

RESULT 62
US-09-495-406-1/c
; Sequence 1, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-0001100S
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-09-495-406-1

Alignment Scores:
Pred. No.: 0.0866 Length: 11474
Score: 110.50 Matches: 78
Percent Similarity: 40.8% Conservative: 44
Best Local Similarity: 26.1% Mismatches: 100
Query Match: 7.1% Indels: 78
DB: 3 Gaps: 17

US-10-734-719-9 (1-291) x US-09-495-406-1 (1-11474)
Qy 35 PheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhe 54
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6218 TTCAAAAAGAAATAGGATGTAATAATACGCTTGCAATTTTACCAGAGATAGATT 6159
Qy 55 PhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu 74
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6158 ATCTTCAAA-----ATAAAA---TTGATTACATCTAAATA-CATCAAAATCAT 6115
Qy 75 LeuIleMetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPh 91
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6114 TTGGTAGTCTTGAATAATCAATTTCTTTTAAACTTGGTCCATTTCCAGCAATAAATAC 6055
Qy 91 eValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys-- 110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6054 T-----TTTCATGATTATTTTCCATCAAGATTATTTTAAAGATTTTATCAAAAT 6004
Qy 111 -----GlnLeuLysGluPheAsnAlaTyr-PheLysPheHisG 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6003 AGATTTTATGTTATAATTAATAATAAGGAATAAATATGTTTAAATAATTTCAA 5944
```

QY 123 luileTyr-----PheAsnGlnArgIleThrSerGlyValTyrMetCysAlaIav 139  
Db 5943 TCATCTTACCAACTTATATATGGAACAATATATAGCAAGGCAATAGAAAGCTGT--- 5888  
QY 139 alalaileAlaLeuGlyTyrLysGluileTyrLeuSerGlyIleAspPheTyrGlnAsnG 159  
Db 5887 -----ATCAATCAGACTTTTAAAGATATAGAAATAATTGTAGTTGAT----- 5846  
QY 159 lySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPhe- 178  
Db 5845 -----GATTGTGGAATGATAATAGTATAGTATAATAGCCAAAGAAATACT 5803  
QY 179 --LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuG 198  
Db 5802 CTAAAAAGACAAAGATAAAATAATCCCAAGTAA-----AAAACTTAG 5755  
QY 198 luPheGluLysThrTyrLys-----IleLysLeuTyrCysLeuCysProAsnSerLeuL 217  
Db 5754 GTCCTTTTAAAGCAAGATATGAAGGTGTGAAGTA-----GCNAACCTCTCCTT 5707  
QY 217 euAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231  
Db 5706 ATATAATCTTTTATAGTCTGATGATTTTGGAACTAAATCTGTGCAAGAGTGTATAA 5647  
QY 232 --IleIleGlnGluLysAsnAsn-----TyrThrLysAspIleLeuLeuProS 247  
Db 5646 AATTTTATAGTAAACAGGATGAAGTTGATTAGTGTGTTTCAATCTATTGTTGAAGTA 5587  
QY 247 erSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe----- 260  
Db 5586 ATGTTTATTCATATAAAAGTTTGACTTTTAAATCTGCTTAAATCTGCTTAAATCTGCTT 5527  
QY 261 --LysLysIleLysIleLysGluAsnValTyrTyr-----LysLeuIleLysA 276  
Db 5526 TAAAAAAATTTATGCAAGAAAAATTTATATTGGACTATGTGGGGAAGTATATAAGAA 5467  
QY 276 sp-----LeuLeuArgLeuProSerAspIleLys 285  
Db 5466 AGAAATGTTATTAAGAGCTTTTGGAGTTTAAAGACTCGAGAAAGATGTTAAA 5414

## RESULT 63

US-09-816-028A-1/c  
; Sequence 1, Application US/09816028A  
; Patent No. 6699705  
; GENERAL INFORMATION:  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/09/816,028A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384  
; OTHER INFORMATION: including LOS biosynthesis locus  
US-09-816-028A-1

## Alignment Scores:

Pred. No.:	0.0866	Length:	11474
Score:	110.50	Matches:	78
Percent Similarity:	40.8%	Conservative:	44
Best Local Similarity:	26.1%	Mismatches:	100

Query Match: 7.1% Indels: 78  
DB: 3 Gaps: 17  
US-10-734-719-9 (1-291) x US-09-816-028A-1 (1-11474)  
QY 35 PheGluAspLysTyrTyrLeuGlyLysCysGlyAlaValPheTyrThrProAsnPhe 54  
Db 6218 TTTCAAAAAAGAAATAGGATTGTAAAAATACGCTTGTCAATTTTACCAGAGATAGTATTT 6159  
QY 55 PhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu 74  
Db 6158 ATCTTCAAA-----ATAAAA-----TTGATTACATCTAAATA-CATCAAAATCAT 6115  
QY 75 LeuIleMetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPh 91  
Db 6114 TTGGTAGTCTGAATTAATCAATTTCTTTTAAACTTGTGTCATTTCCAGCAATAATACTT 6055  
QY 91 eVallysthrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys-- 110  
Db 6054 T-----TTTCATGATTATTTTCCATCAAGATTATTTTAAACGATTATTTATCAAAAT 6004  
QY 111 -----GlnLeuLysGluPheAsnAlaTyr-PhelysPheHisG 123  
Db 6003 AGATTTTATGTTATTAATAATAATAAATAAGCAATAATAATATGTTTAAAAATTTCAA 5944  
QY 123 luileTyr-----PheAsnGlnArgIleThrSerGlyValTyrMetCysAlaV 139  
Db 5943 TCATCTTACCAACTTATATATGGAACAATATATAGCAAGGCAATAGAAAGCTGT--- 5888  
QY 139 alalaileAlaLeuGlyTyrLysGluileTyrLeuSerGlyIleAspPheTyrGlnAsnG 159  
Db 5887 -----ATCAATCAGACTTTTAAAGATATAGAAATAATTGTAGTTGAT----- 5846  
QY 159 lySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPhe- 178  
Db 5845 -----GATTGTGGAATGATAATAGTATAGTATAATAATATAGCCAAAGAAATACT 5803  
QY 179 --LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuG 198  
Db 5802 CTAAAAAGACAAAGATAAAATAATCCCAAGTAA-----AAAACTTAG 5755  
QY 198 luPheGluLysThrTyrLys-----IleLysLeuTyrCysLeuCysProAsnSerLeuL 217  
Db 5754 GTCCTTTTAAAGCAAGATATGAAGGTGTGAAGTA-----GCNAACCTCTCCTT 5707  
QY 217 euAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231  
Db 5706 ATATAATGTTTATAGTCTGATGATTTTGGAACTAAATCTGTGCAAGAGTGTATAA 5647  
QY 232 --IleIleGlnGluLysAsnAsn-----TyrThrLysAspIleLeuLeuProS 247  
Db 5646 AATTTTATAGTAAACAGGATGAAGTTGATTAGTGTGTTTCAATCTATTGTTGAAGTA 5587  
QY 247 erSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe----- 260  
Db 5586 ATGTTTATTCATATAAAAGTTTGACTTTTAAATCTGCTTAAATCTGCTTAAATCTGCTT 5527  
QY 261 --LysLysIleLysIleLysGluAsnValTyrTyr-----LysLeuIleLysA 276  
Db 5526 TAAAAAAATTTATGCAAGAAAAATTTATATTGGACTATGTGGGGAAGTATATAAGAA 5467  
QY 276 sp-----LeuLeuArgLeuProSerAspIleLys 285  
Db 5466 AGAAATGTTATTAAGAGCTTTTGGAGTTTAAAGACTCGAGAAAGATGTTAAA 5414

## RESULT 64

US-10-303-162-1/c  
; Sequence 1, Application US/10303162  
; Patent No. 6723545  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada

```

, TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
,
, TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
,
, FILE REFERENCE: 019633-000111US
,
, CURRENT APPLICATION NUMBER: US/10/303,162
,
, CURRENT FILING DATE: 2002-11-21
,
, PRIOR APPLICATION NUMBER: US/09/816,028
,
, PRIOR FILING DATE: 2001-03-21
,
, PRIOR APPLICATION NUMBER: US 60/118,213
,
, PRIOR FILING DATE: 1999-02-01
,
, PRIOR APPLICATION NUMBER: US 09/495,406
,
, PRIOR FILING DATE: 2000-01-31
,
, NUMBER OF SEQ ID NOS: 49
,
, SOFTWARE: PatentIn Ver. 2.1
,
, SEQ. ID. NO 1
,
, LENGTH: 11474
,
, TYPE: DNA
,
, ORGANISM: Campylobacter jejuni
,
, FEATURE:
,
, OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
,
, OTHER INFORMATION: including LOS biosynthesis locus
,
, US-10-303-162-1

```

Alignment Scores:			
Pred. No.:	0.0866	Length:	11474
Score:	110.50	Matches:	78
Percent Similarity:	40.8%	Conservative:	44
Best Local Similarity:	26.1%	Mismatches:	100
Query Match:	7.1%	Indels:	78
DB:	3	Gaps:	17
US-10-734-719-9 (1-291) x US-10-303-162-1 (1-11474)			
Qy	35	PheGluaspIysTyrTyrrLeuGlyLysCysLysAlaValPheTyrThrProAsnPhe	54
Db	6218	TTCAAAAAGAAATAGGATGTAAATACTGCTTGCATTTTTTACCAAGATAGTATTT	6159
Qy	55	PhePheGluGlnTyrTyrrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu	74
Db	6158	ATCTTCAA-----ATAAAA--TTGATTACATCTAAATA-CATCAAAATCAT	6115
Qy	75	LeuIleMetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPh	91
Db	6114	TTGGTAGTCTTGAATAATCAATTTCTTTTAAACTTGCTCCATTTCCAGCAATATAACTT	6055
Qy	91	eValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys--	110
Db	6054	T-----TTTCATGATTATTTTCATCAAGATTATTTTAAACGATTTTATCAAAAT	6004
Qy	111	-----GlnLeuLysGluPheAsnAlaTyr-PheLysPheHisG	123
Db	6003	AGATTTTATGTTATAATTAAATTAATAAATAAGGAATAAATATGTTTAAATTTCAA	5944
Qy	123	luIleTyr-----PheAsnGlnArgIleThrSerGlyValTyrMetCysAlav	139
Db	5943	TCATCTTACCAACTTATAATGTGGAACAATATATAGCAAGGCAATAGAAAGCTGT---	5888
Qy	139	alalalleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr-GlnAsnG	159
Db	5887	-----ATCAATACGACTTTTAAAGATATAGAAATAATGTAGTTGAT-----	5846
Qy	159	lySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPhe-	178
Db	5845	-----GATGTGGAATGCATAATAGTATAAATATATAGCCAAAGAATACT	5803
Qy	179	--LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuG	198
Db	5802	CTAAAAAGACAAAAGAAATAAAAAATAATCCACAAATGAA-----AAAAACTTAG	5755
Qy	198	luPheLeuGluLysThrTyrLys---lleLysLeuTyrCysLeuCysProAsnSerLeuL	217
Db	5754	GTCTTTTAAAGCGAAGATATGAGGTTGGAAGTA-----GCAAACTCTCCTT	5707
Qy	217	euAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe-----	231

Db	5706	ATATAAGTTTITAGATCCTGATGATTATTITGGAACTAAATAGCTTGTGAAGAGTGTTATAA	5647
Qy	232	--llelleGlnGluLysAsnAsn-----TyrThrLysAspilleLeullePros	247
Db	5646	AAATTTAGATGAACAGGATGAAGTTGATTTAGTGTTTTTCAATGCTATTGTTGAAAGTA	5587
Qy	247	erSerGluAlaTyrGlyLysPheSerLysAsnilleAsnPhe-----	260
Db	5586	ATGTTATTTTCATATAAAAGTTTGACATTTAATCTCGTTTTTATATAGCAAAAAGAGTTTG	5527
Qy	261	--LysLysilleLysilleLysGluAsnValTyrTyr-----LysLeulleLysA	276
Db	5526	TAAAAAAAATTATTGCAAGAAAGAAATTTATTATTTGGACTATGTGGGGAAACTTATAGAA	5467
Qy	276	sp-----LeuLeuArgLeuLeuProSerAspileLys	285
Db	5466	AGAAATGTATTAGAACGTTTTTCGAGGTTTAAGACTCGAGAAAGATGTTAAA	5414
RESULT 65			
US-10-303-134-1/c			
; Sequence 1, Application US/10303134			
; Patent No. 6825019			
; GENERAL INFORMATION:			
; APPLICANT: Gilbert, Michel			
; APPLICANT: Wakarchuk, Warren W.			
; APPLICANT: National Research Council of Canada			
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of			
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics			
; FILE REFERENCE: 019633-000111US			
; CURRENT APPLICATION NUMBER: US/10/303,134			
; PRIORITY FILING DATE: 2002-11-21			
; PRIOR APPLICATION NUMBER: US/09/816,028			
; PRIOR FILING DATE: 2001-03-21			
; PRIOR APPLICATION NUMBER: US 60/118,213			
; PRIOR FILING DATE: 1999-02-01			
; PRIOR APPLICATION NUMBER: US 09/495,406			
; PRIOR FILING DATE: 2000-01-31			
; NUMBER OF SEQ ID NOS: 49			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 11474			
; TYPE: DNA			
; ORGANISM: Campylobacter jejuni			
; FEATURE:			
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384			
US-10-303-134-1			
Alignment Scores:			
Pred. No. : 0.0866 Length: 11474			
Score: 110.50 Matches: 78			
Percent Similarity: 40.8% Conservative: 44			
Best Local Similarity: 26.1% Mismatches: 100			
Query Match: 7.1% Indels: 78			
DB: 3 Gaps: 17			
US-10-734-719-9 (1-291) x US-10-303-134-1 (1-11474)			
Qy	35	PheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhe	54
Db	6218	TTCAAAAAAAGATAGGATTGTAATAATFACTGCCCTTGCAATTTTCCACAGATAGTATT	6155
Qy	55	PhePheGluGlnTyrTyrTyrLeuLysHisLeulleGlnAsnGlnGluTyrGluThrGlu	74
Db	6158	ATCTTCAAA-----ATAAAA---TTGATTACATCTAATA-CATCAAAATCAT	6111
Qy	75	LeulleMetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPh	91
Db	6114	TTGGTAGCTTTGAATAATCAATTTCTTTTAAACTTGGTCCATTTCCACCAATAATAACTT	6055
Qy	91	eValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys--	110

Db 6054 T-----TTTCATGATTATTTTCATCAAGATTATTTATTTTAAACGATTATTTATCAAAAT 6004  
Qy 111 -----GlnLeuLysGluPheAsnAlaTyr-PheLysPheHisG 123  
Db 6003 AGATTTTATGTTATAATAATATAATAATAAAGGAATAAATATGTTTAAATTTCAA 5944  
Qy 123 lullletyr-----PheAsnGlnArgilleThrSerGlyValTyrMetCysAlav 139  
Db 5943 TCATCTTACCAACTTATAATGTCGAACAATATATAGCAAGGCAATAGAAAGCTGT--- 5888  
Qy 139 alAlalleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnG 159  
Db 5887 -----ATCAATCAGACTTTTAAAGATATAGAAATAATTTAGTTGAT----- 5846  
Qy 159 lySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLeuAlaProAspPhe- 178  
Db 5845 -----GATTTGGAAATGATAATATATAGTATAATAATATAGCCAAAGAACTACT 5803  
Qy 179 --LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuG 198  
Db 5802 CTAAAAAGCAAAAGAAATAATAATATCCACATGAA-----AAAACTTAG 5755  
Qy 198 luPheLeuGluLysThrTyrLys---lleLysLeuTyrCysLeuProAsnSerLeuL 217  
Db 5754 GTCCTTTTAAAGCAAGATATGAAGGTGTGAAAGTA-----GCAAACTCTCCTT 5707  
Qy 217 euAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231  
Db 5706 ATATAATGTTTATAGATCTCTGATGATTTTGAACCTAAATGCTTGTGAAGAGTGATAA 5647  
Qy 232 --llelleGlnGluLysAsnAsn-----TyrThrLysAspIleLeuLeuProS 247  
Db 5646 AAATTTTATGATCAACAGCATGAAGTTGATTTAGTTTCAATGCTATTTGTTGAAAGTA 5467  
Qy 276 sp-----LeuLeuArgLeuProSerAspIleLys 285  
Db 5466 AGAAATTTGATTTTAAAGCTTTTCGAGTTTAAAGCTCGAGAAAGATGTTAAA 5414

## RESULT 66

US-10-303-118-1/c  
; Sequence 1, Application US/10303118  
; Patent No. 6905867  
; GENERAL INFORMATION:  
; APPLICANT: Gilbert, Michel  
; APPLICANT: Wakarchuk, Warren W.  
; APPLICANT: National Research Council of Canada  
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of  
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics  
; FILE REFERENCE: 019633-000111US  
; CURRENT APPLICATION NUMBER: US/10/303,118  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US/09/816,028  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/118,213  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: US 09/495,406  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 11474  
; TYPE: DNA  
; ORGANISM: Campylobacter jejuni  
; FEATURE:  
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384

; OTHER INFORMATION: including LOS biosynthesis locus  
US-10-303-118-1

## Alignment Scores:

Pred. No.: 0.0866 Length: 11474  
Score: 110.50 Matches: 78  
Percent Similarity: 40.8% Conservative: 44  
Best Local Similarity: 26.1% Mismatches: 100  
Query Match: 7.1% Indels: 78  
DB: 3 Gaps: 17

US-10-734-719-9 (1-291) x US-10-303-118-1 (1-11474)

Qy 35 PheGluAspLysTyrTyrLeuGlyLysCysLysAlaValPheTyrThrProAsnPhe 54  
Db 6218 TTCAAAAAAAGAAATAGGATTGTAATAATCTGCTTGCATTTTACCAAGATAGTATTT 6159  
Qy 55 PhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGlu 74  
Db 6158 ATCTTCAA-----ATAAAA---TTGATTACATCTAAATA-CATCAAAATCAT 6115  
Qy 75 LeuLeuMetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPhe 91  
Db 6114 TTGTTAGTCTTCAATAATAATCTTTTAAACTTGGTCCATTTCCAGCAATAATAACTT 6055  
Qy 91 eValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys-- 110  
Db 6054 T-----TTTCATGATTAATTTTCCATCAAGATTATTTTAAAGATTTTATCAAAAT 6004  
Qy 111 -----GlnLeuLysGluPheAsnAlaTyr-PheLysPheHisG 123  
Db 6003 AGATTTTATGTTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5944  
Qy 123 luilleTyr-----PheAsnGlnArgilleThrSerGlyValTyrMetCysAlav 139  
Db 5943 TCATCTTACCAACTTATAATGTCGAACAATATATAGCAAGGCAATAGAAAGCTGT--- 5888  
Qy 139 alAlalleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnG 159  
Db 5887 -----ATCAATCAGACTTTTAAAGATATAGAAATAATTTAGTTGAT----- 5846  
Qy 159 lySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPhe- 178  
Db 5845 -----GATTTGGAAATGATAATAGTATAATAATAATAATAATAATAATAATA 5803  
Qy 179 --LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuG 198  
Db 5802 CTAAAAAGCAAAAGAAATAATAATCCACATGAA-----AAAACTTAG 5755  
Qy 198 luPheLeuGluLysThrTyrLys---lleLysLeuTyrCysLeuProAsnSerLeuL 217  
Db 5754 GTCCTTTTAAAGCAAGATATGAAGGTGTGAAAGTA-----GCAAACTCTCCTT 5707  
Qy 217 euAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231  
Db 5706 ATATAATGTTTATAGATCTCTGATGATTTTGAACCTAAATGCTTGTGAAGAGTGATAA 5647  
Qy 232 --llelleGlnGluLysAsnAsn-----TyrThrLysAspIleLeuLeuProS 247  
Db 5646 AAATTTTATGATCAACAGCATGAAGTTGATTTAGTTTCAATGCTATTTGTTGAAAGTA 5587  
Qy 247 erSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe----- 260  
Db 5586 ATGTTTATTTTCATATAAAAAAGTTTGACTTTTAAATCTGCTATTTGTTGAAAGTA 5467  
Qy 276 sp-----LeuLeuArgLeuProSerAspIleLys 285  
Db 5466 AGAAATTTGATTTTAAAGCTTTTCGAGTTTAAAGCTCGAGAAAGATGTTAAA 5414





6158 ATCTTCAA--ATAAAA--TTGATTACATCTAAATA-CATCAAAATCAT 6115

75 LeuileMetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu-AsnGluAsnPh 91

6114 TTGGTAGCTTGAATAAATCAATTTCTTTAAACTTGGTCCATTTCCAGCAATAATAACTT 6055

91 eValLysThrPheTyrAspTyrPheProaspAlaHisLeuGlyTyrAspPheLys-- 110

6054 T-----TTTCATGATTATTTTCCATCAAGATTATTTTAAACGATTATTCACAAAT 6004

111 -----GlnLeuLysGluPheAsnAlaTyr--PheLysPheHisG 123

6003 AGATTTTATGTTATAATTAATAATAATAAAGAAATAAATATATGTTTAAAAATTTCAA 5944

123 luileTyr-----PheAsnGlnArgileThrSerGlyValTyrMetCysAlav 139

5943 TCACTTTACCAACTTATAATGTGGCAATATATAGCAAGGCAATAGAAAGCTGT---- 5888

139 alAlaileAlaLeuGlyTyrLysGluileTyrLeuSerGlyIleAspPheTyrGlnAsnG 159

5887 ----ATCAATCAGACTTTTAAAGATATAGAAAATAATTGTAGTTGAT----- 5846

159 lySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProaspPhe- 178

5845 -----GATTCGGAAATGATATAATAGTATAATAATATAGCCAAAGAACTACT 5803

179 --LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuG 198

5802 CTAATAAGACAAAAGATAAAAATAATCCACAATGNA-----AAAACTTAG 5755

198 luPheLeuGluLysThrTyrLys--IleLysLeuTyrCysLeuCysProAsnSerLeuL 217

5754 GTCCTTTAAGACGACGATATGAAGGTGTGAAGTA-----GCAAACTCTCCTT 5707

217 euAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231

5706 ATATAATCTTTTATGATCTCGATGATTATTGTGAACATAATGCTGTGAAGAGTGTATAA 5647

232 --IleIleGlnGluLysAsnAsn-----TyrThrLysAspIleLeuileProS 247

5646 AAAATTTAGATGAACAGAGTAAGTGTGATTTAGTGTGTTTCAATGCTATGTTGNAAGTA 5587

247 erSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe----- 260

5586 ATGTTATTTCATATAAAAAGTTTGACTTTTAATCTGTTTTTATAGCAAAAAGAGTTTG 5527

261 --LysLysIleLysIleLysGluAsnValTyr-----LysLeuileLysA 276

5526 TAAAAAAAATTTATGCAAGAAAAATTTATTTATGACTATGTGGGGGAAACTTATAAGAA 5467

276 sp-----LeuLeuArgLeuProSerAspIleLys 285

5466 AGAAATTGATTATAGAACGCTTTTGGAGTTTATAGACTCGAGAAAGATGTTAAA 5414

RESULT 69

US-09-662-254B-23

; Sequence 23, Application US/09662254B

; Patent No. 6933145

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Li, Yi

; TITLE OF INVENTION: Bawden, Alison Louise

; TITLE OF INVENTION: Vertebrate Cells

; FILE REFERENCE: UF-221C1XC1

; CURRENT APPLICATION NUMBER: US/09/662,254B

; CURRENT FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 09/086,651

; PRIOR FILING DATE: 1998-05-29

; PRIOR APPLICATION NUMBER: 60/224,479

; PRIOR FILING DATE: 2000-08-10

; NUMBER OF SEQ ID NOS: 80

; SOFTWARE: PatentIn version 3.1

Sequence of Heterologous

```
Db 48294 TCACGATTAACGAGTACCTTTAGTATCATCAATTAATGCAAA----- 48339
Qy 235 uLysAsnAsnTyrThrLysAspLeuLeuProSerSerGluAlaTyrGlyLysPheSe 255
Db 48340 -----GGATCCATGTTAATACAACTTCCCAATATATTTTTTTTGC 48383
Qy 255 rLysAsnLe-----AsnPheLysLysLysLysLysLysLysGluAsnVa 269
Db 48384 TAATAATATCATCAATATACATCTTCTCATATTCAAATTCAAATATATCTGTTAATAT 48443
Qy 269 lTyrTyrLysLeuLysAspLeuLeuArgLeuProSerAspLysHis 286
Db 48444 TTTATTTTCTTA-----TCATTTATAACATTACCTTAAATAATGTAACACAT 48489

RESULT 70
US-08-545-528D-1
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment
; FILE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB193PI
; CURRENT APPLICATION NUMBER: US/08/545,528D
; CURRENT FILING DATE: 1995-10-19
; PRIOR APPLICATION NUMBER: US 08/488,018
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/473,545
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Mycoplasma genitalium
US-08-545-528D-1

Alignment Scores:
Pred. No.: 26.6 Length: 580073
Score: 110.00 Matches: 46
Percent Similarity: 37.6% Conservative: 37
Best Local Similarity: 20.8% Mismatches: 58
Query Match: 7.0% Indels: 80
DB: 3 Gaps: 11

US-10-734-719-9 (1-291) x US-08-545-528D-1 (1-580073)
Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGlu----- 87
Db 77311 TGCAGTGATTATTCCTATAATCATGTTGATTATTTCTTTGTTTCAACTACCTTAATTC 77370
Qy 88 -----AsnGlu 89
Db 77371 AGAGTTGAAAAAATTTGCTATTAGATTAAAGCATTTGGGATTTCCAAATTTAAATAATCT 77430
Qy 90 AsnPheVal-----LysThrPheTyrAspTyrPhePro 100
Db 77431 CGCTTCATTTTATCAATATACATCCCTTTATTTGGCCCTTTGTTGTTGATTTCATCC 77490
Qy 101 AspAlaHisLeuGlyTyr-----AspPhePheLysGlnLeuLysGlu 114
Db 77491 CTTTTCATCTATCTAATTTGCATACATAATGAGGTAAATTTTGCAAGCTCATCGATCTT 77550
Qy 115 PheAsnAlaTyrPheLysPheHisGluLe-----TyrPheAsnGlnArgIleThrSerGly 133
Db 77551 TTTAGATGCTTTTAAAGTTTGAAGTGCATGTTGTTCAATGTAGTTTACTAGCGGT 77610
Qy 134 Val-----TyrMetCysAlaValAlaLeu-----Gly 144
Db 77611 TTTATCAATACCTTTGTTGTAATGATTAGATTGAACAAATAAGATTGACAAGGA 77670
```

```
Qy 145 TyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer----- 160
Db 77671 AATCAAAACTCCTAAG-----GATTTTCTTTTAAACAAATCATAAACCC 77721
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 77722 AACCAGAAATTCACCTAGTAATGAAGCTGAATTTCTACAGATAGCTACTGATTATTGGAG 77781
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 77782 GAAAGTCAAAATATCTCAA-----AAGGGTTTAAAGCAATTA 77820
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 77821 AAAAAAGAATATAA-----AGATCCATTATTATAAC--- 77853
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 77854 -----CCTAACCTTGAATATAAACCGTTTGTAAATGAAAAAGAAATTTCACT 77901

RESULT 71
US-09-662-254B-32
; Sequence 32, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; FILE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UF-221C1XC1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(2469)
; OTHER INFORMATION:
US-09-662-254B-32

Alignment Scores:
Pred. No.: 0.019 Length: 2469
Score: 108.00 Matches: 80
Percent Similarity: 32.2% Conservative: 47
Best Local Similarity: 20.3% Mismatches: 124
Query Match: 6.9% Indels: 144
DB: 3 Gaps: 18

US-10-734-719-9 (1-291) x US-09-662-254B-32 (1-2469)
Qy 19 SerArgLeuProAsnAspPheAspVal---PheArgCysAsnGlnPheTyrPheGluAsp 37
Db 1291 TCTAGATAATATTATTGATTTTTTTTATTGATATTAAACAAATAGATTAGATATCAAGAT 1350
Qy 38 LysTyrTyrLeuGlyLysLysCysLysAlaValPheTyr-----ThrProAsnPhe 54
Db 1351 AAATATAAAAAACAAATCTCTAACTCCAAATTTATTTATAGATTGTCAATAATAATT 1410
Qy 55 PhePhe-----GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsn 68
```

```
Db 1411 TTTATAGCAGTTTATAATGAAAAAGAACAAATATGCCGAAGAAACAACTAAACATGTTT 1470
Qy 69 GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGln-----AlaHisLeu 86
Db 1471 ATATATTCGGAATATCTTTATTTATTAATTAAGTAAATTTAATGAATTAATAGGTATATA 1530
Qy 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 1531 AAAAATAATAAAAAATTAACAACTATATTTGATTAATCAAAATGATATTAATAATTA 1590
Qy 107 AspPhePhe-----Lys 110
Db 1591 GATAATTTTAAAAAGACTGCTGTTTCATATATATAAGTAGGAATAGATTAAATAGATAAA 1650
Qy 111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130
Db 1651 AAAAGTAGAGAAATGATTAATATGATGACTATAATGATGTTTATTTAAAT---ATATTA 1707
Qy 131 ThrSerGlyValTyrMetCys----- 137
Db 1708 ACTCCGAAATTAATAATCGTGTATATAATATAATTAATTAATAGATTATATAAAATATAGAT 1767
Qy 138 -----AlaValAlaIleAlaLeuGly 144
Db 1768 ATTTTAAATATGATTATATAGAAATTAACAGATATTCATTACTACCGTAAACATTAGGA 1827
Qy 145 TyrLysGlu----- 147
Db 1828 TATAAACACAAAAATATTGATACCTGCTCTACAATATCTTTTACCACCTCGAAGATGTA 1887
Qy 148 IleTyrLeuSerGlyIleAspPheTyrClnAsnGlySerSerTyrAla----- 163
Db 1888 ATTAATTTATAATAATGTAATATTTTCTGAATATTTACTATAGAAATATATTAGATATATTACATACGATACG 1947
Qy 164 -----PheAspThrLysGlnGluAsnLeu-----LeuLysLeu 174
Db 1948 TTAAAAATTAATAATTTATCTGATTTTGATATTAAAGATATAAATGTTGAATTAATAACT 2007
Qy 175 AlaProAspPheLysAsnAspArgSerHisTyr----- 185
Db 2008 ATAAATCGAAAGATTAATTTCTGAATATTTACTATAGAAATATTAGTATATAATACTTTGAA 2067
Qy 186 -----IleGlyHis----- 188
Db 2068 CAGATGGATAATTATAATTTTATATAGATATAGGACAAAAATATTTTATATATAAAT 2127
Qy 189 -----SerLysAsnThrAspIleLys----- 195
Db 2128 GATGTATATCGAATAGTAATATTGTAATAATAAAGTAATATTATTCTAAATAATGNAAT 2187
Qy 196 -----AlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCys 212
Db 2188 TTTGGTGATCTTTGGCAATTTTAAATAAATAATATAAATTTCTATTATACATTATTTT 2247
Qy 213 ProAsn---SerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231
Db 2248 GATAATCTGAATTTATTAATAAATTTT-----TTATATCCGAATGTTTAAATAATATATTT 2301
Qy 232 IleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr 251
Db 2302 -----AATTATGATCAAGATTATATACTAGATGATTATTTTCATTAT 2343
Qy 252 -----GlyLysPheSerLysAsnIleAsnPheLysLysLys 264
Db 2344 ATGTTTATAATATATTAATTTTCATTATTAATACTAATAATA---TTATCATGGATAGAT 2400
Qy 265 IleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArg 279
Db 2401 GTAAACAAAGATAATAATCTAAATTTATATGATAATACTTTTAAGA 2445
```

RESULT 72

US-09-662-254B-24

; Sequence 24, Application US/09662254B

; Patent No. 6933145

; GENERAL INFORMATION:

; APPLICANT: Moyer, Richard W.

; APPLICANT: Li, Yi

; APPLICANT: Baden, Alison Louise

; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous

; FILE REFERENCE: UF-221C1XC1

; CURRENT APPLICATION NUMBER: US/09/662,254B

; CURRENT FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: 09/086,651

; PRIOR FILING DATE: 1998-05-29

; PRIOR APPLICATION NUMBER: 60/224,479

; PRIOR FILING DATE: 2000-08-10

; NUMBER OF SEQ ID NOS: 80

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 24

; LENGTH: 50000

; TYPE: DNA

; ORGANISM: Amsacta moorei entomopoxvirus

; US-09-662-254B-24

Alignment Scores:

Pred. No.: 1.38 Length: 50000

Score: 108.00 Matches: 80

Percent Similarity: 32.2% Conservative: 47

Best Local Similarity: 20.3% Mismatches: 124

Query Match: 6.9% Indels: 144

DB: 3 Gaps: 18

US-10-734-719-9 (1-291) x US-09-662-254B-24 (1-50000)

Qy 19 SerArgLeuProAsnAspPheAspVal---PheArgCysAsnGlnPheTyrPheGluAsp 37

Db 5588 TCTAGATAAATATTGATTTTTTTTATTTGATATTAAACAAATAGTAGAATATCAAGAT 5647

Qy 38 LysTyrTyrLeuGlyLysLysCysLysAlaValPheTyr-----ThrProAsnPhe 54

Db 5648 AAATATAAAACAAATCTCTAACTCCAAATTTATTTTATAAGATTGTCAAATAATTTA 5707

Qy 55 PhePhe-----GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsn 68

Db 5708 TTTATAGCAGTTTATAATGAAAGAACAAATATGCGGAAGAAAGACAACTAAACATGTTT 5767

Qy 69 GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGln-----AlaHisLeu 86

Db 5768 ATATATTTCGGAATATCTTTTATTTATTTAAGTAATTTTAAATGAATTAATAGGTATATA 5827

Qy 87 GluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106

Db 5828 AAAATAATAAAAAATTAACAACTATATTTTGATTAATCAAAATGATATAAAATAATTTA 5887

Qy 107 AspPhePhe-----Lys 110

Db 5888 GATAATTTTATAAAGATACGTATTCATATATAAGTAGGAATAGATTAAATAGATAAA 5947

Qy 111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130

Db 5948 AAAAGTAGAGAAATGATTAATTAATGATGACTATAATTTGATGTTTATTTAAAT---ATATTA 6004

Qy 131 ThrSerGlyValTyrMetCys----- 137

Db 6005 ACTCCGAAATTAATAATCGTGTATATAATTAATTAATAGATTATATAAAATATAGAT 6064

Qy 138 -----AlaValAlaIleAlaLeuGly 144

Db 6065 ATTTTAAATATGATTATATAGAAATTAACAGATATTTCCATCTACTACCGTAAACATTAGGA 6124

Qy 145 TyrLysGlu----- 147

Db 6125 TATAAACACAAAAATATTGATGACTGCTCTACAATATCTTTTTTACCACCTCGAAGATGTA 6184

Qy 148 IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAla----- 163



```
Qy 202 LysThrTyrLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhelle 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 649 CAACAATTCATTGTAATAATATACAGCTGAGCAAGAT-----TTTCCAAAATTATC 702
Qy 222 GluLeuAlaProAsnLeuAsnSerAsnPhellelleGlnGluLysAsnAsnTyrThrLys 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 703 GGTTCTCCACCA-----ATCGATTATTGATCAGCAAAAAAATTTCTTACGCTAAA 753
Qy 242 AspileLeuLeuProSerGluAlaTyrGlyLysPheSerLysAsnLeuAsnPhelys 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 754 GACCTTCTT-----CGGTATACAGAAATTTACGGTTCGAAATGATTTCCGAA 798
Qy 262 LysileLysileLysGluAsnValTyrTyr---LysLeuLeuLysAspLeuLeuArgLeu 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 799 GAAATAGGATTGAGATTTTGCTTACTTTAGCCGCTGTTCAAGCAAAACAGGACTA 858
Qy 281 ---ProSerAspileLysHisTyr 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 859 TCGCCGAGAGACTTTCGTAAATTC 882

RESULT 74
US-09-601-198-18
; Sequence 18, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-18

Alignment Scores:
Pred. No.: 0.00761 Length: 1182
Score: 107.50 Matches: 62
Percent Similarity: 36.1% Conservative: 39
Best Local Similarity: 22.1% Mismatches: 85
Query Match: 6.9% Indels: 94
DB: 3 Gaps: 12

US-10-734-719-9 (1-291) x US-09-601-198-18 (1-1182)
Qy 59 TyrTyrThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGlu---ThrGluLeuLeuMet 77
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 40 TATTTTACATTTGGTATTATTTATTAATGATGTAACATAAAGTTAATGATTAATTTT 99
Qy 78 CysSer-----AsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPhelysThr 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 ATAAAAATCTATTCACAAACAATAATATTAATGATCATCAAAATTTTTCGAAACA 159
Qy 95 PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 -----GATTTAAAAATCCAACTTAATACA 183
Qy 115 PheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyVal 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 184 TTGATTAACCTGTTTAAA---AAAGAATTGCACATTAAT----- 219
Qy 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154
```

```
Db 220 TTTATTTGATGAATTTATTTAGTTTATGTTTCTTAACTACACATACTACT 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeu 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 TTTTAT-----AACTATAATGAATTTATGAATTA 309
Qy 175 AlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspile 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 AGATCAAAATTTTAATAATGACACAGTGAATTATTAAAGTTTCAGATCAAAAAGTTATATA 369
Qy 195 LysAlaLeuGluPhe-----LeuGluLysThrTyrLysIleLysLeu 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 ---GCAGTTGATTTTGAATAAGCCACTATTCTAAAGAAAAATAGTTATAAGTTGATTTCA 426
Qy 209 Tyr-----CysLeuCysProAsnSer 215
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 427 TATTATATTTTGAATATGATTATCAATTAATAATTGATTTGTGCCAATCATTTAATTTA 486
Qy 216 LeuLeuAlaAsnPhelleGluLeuAlaProAsnLeuAsn----- 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 487 AAAATTAAGCAATTTATAGTTTCACAAGCTCATTTAAACGGTTTAAATTTGATAATGAA 546
Qy 228 ----- 228
Db 547 GATAATAAACTTCATTATTATTAGAAATTAGAAAAATTTTAAAAATAAATTTTAACTAT 606
Qy 229 -----SerAsnPhellelleGlnGluLysAsnAsnTyr---ThrLys 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 607 AATGAATATAATTTAATTAGTAATGTATTAGTGAGACAAAAACCACTTTTGCCTGGCT 666
Qy 242 AspileLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnLeuAsnPhelys 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 667 GATATCATTTACTCAAAACAGCACAA-----TCATTAATATAATTAATCCTAAA 711
Qy 262 LysileLysileLysGluAsn-ValTyrTyrLysLeuLeuLysAspileLeuLeuArgLeu 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 712 AAGATTGAGAGCTACTTTTAAATGATATAATAATAATAAAGTTATTGAATTA 769

RESULT 75
US-09-662-254B-23/c
; Sequence 23, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; FILE REFERENCE: UP-221C1XCI
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 23
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Anasacta moorei entomopoxvirus
US-09-662-254B-23

Alignment Scores:
Pred. No.: 1.58 Length: 50000
Score: 107.50 Matches: 71
Percent Similarity: 37.0% Conservative: 47
Best Local Similarity: 22.3% Mismatches: 100
Query Match: 6.9% Indels: 102
DB: 3 Gaps: 16

US-10-734-719-9 (1-291) x US-09-662-254B-23 (1-50000)
```





```

; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (65..1459)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1474..2151
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2239..2475)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2502..2987
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3080..6091
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (6277..6768)
;
US-08-107-755A-1
Alignment Scores:
Pred. No.: 0.157 Length: 6768
Score: 105.50 Matches: 56
Percent Similarity: 39.7% Conservative: 37
Best Local Similarity: 23.9% Mismatches: 82
Query Match: 6.7% Indels: 59
DB: 2 Gaps: 12

US-10-734-719-9 (1-291) x US-08-107-755A-1 (1-6768)
Qy 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCys 45
Db 691 GATCATTTTACTTGTTCGCGTAATA-----GATAAACACITT---GATAAGATAAA 644
Qy 46 LysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
Db 643 AAAGTCGCATAC-----TTTTTAAATAGTAGTGGTTATATATACCAGAACTTATA 596
Qy 66 IleGluAsnGlnGluTyr-----GluThrGluLeuIleMetCysSer----- 79
Db 595 AAACAAATAAAAAATATATGTTTATTGAATCTGATATGACTATATAAAAGCCATAAACAC 536
Qy 80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
Db 535 TACAATAGTACTCTTAATACTAATATTGCTTATTATACATTGATGACTATCAGAAATAT 476
Qy 88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 475 TTAATGATATATTAAAAATGTAATATTACTTTTAACTTTTAACTTTTGAATTTACAATAT 416
Qy 107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
Db 415 GATAGTCCCGATTGTGTATGTTTAAATATAATATTTTATATATATAGTTTATTTCAC 356
Qy 117 -----AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal 134
Db 355 ATAAATCTAAATTTGAATTTAAAAAATTAATATTATTCTATGAGTTTATTGGTGATTA 296
Qy 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154
Db 295 TTGGCTAGTAGTTATAGAGGTGCATTATTATTCTAGATATGATATTAATAGTATTGAC 236
Qy 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeu 174
Db 235 GAATTTAAAAAT---ACATTAGAAATATTTAATATAAAAAATAAAAAAGTTTATGGAATTA 179
Qy 175 -----AlaProAspPheLysAsn 180
Db 178 ATAGATATGTATAAAAAAATAGTAATAGAAATTTATGAATGTATGTATGTAATAAATAAAT 119
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu 197
Db 118 GATTATGATAGTTATATTGATATGATGAGAGAAATAGCTTTAGAACTCTAATATTTAATATCTT 59
Qy 198 GluPhe-----LeuGluLysThrTyrLysIleLysLeuTyr 209
Db 58 GAATTTTAAATATCAATTAATTAATCATACAAATTTGTACTATAT 17

RESULT 79
US-07-991-867B-1/c
; Sequence 1, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8457 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (65..1459)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1474..2151
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2239..2475)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2502..2987
; NAME/KEY: CDS
```



```

; LOCATION: 3080..6091
; FEATURE: CDS
; NAME/KEY: complement (6277..6768)
; LOCATION: complement (6277..6768)
US-07-991-867B-1
Alignment Scores:
Pred. No.: 0.215 Length: 8457
Score: 105.50 Matches: 56
Percent Similarity: 39.7% Conservativity: 37
Best Local Similarity: 23.9% Mismatches: 82
Query Match: 6.7% Indels: 59
DB: 2 Gaps: 12

US-10-734-719-9 (1-291) x US-07-991-867B-1 (1-8457)
Qy 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCys 45
Db 691 GATCATTTTCTGTTGGTAATA-----GATAAACACTTT--GATAAAGATAAA 644
Qy 46 LysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
Db 643 AAAGCTGCATAC-----TTTTTAATAGTAGTGGTTATATACCAGAACTTATA 596
Qy 66 IleGlnAsnGlnGluTyr-----GluThrGluLeuIleMetCysSer----- 79
Db 595 AAACAAATAAAATAATATATGTTTATGAACTCTGATATGACTATATAAAGCCATAAACAC 536
Qy 80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
Db 535 TACAATAGTACTCTCTAATACTAATATGCTTTATTTATACATGATGTACTATCAGAATAT 476
Qy 88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 475 TTAATGATATATTTAAATGTAATATTTACTTTTTTAACTATTTTGAATTTACAATAT 416
Qy 107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
Db 415 GATAGTCCCGATGTTGGTATGTTTAAATAATAATATTTTATATATATATAGTTTATTTCAC 356
Qy 117 -----AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal 134
Db 355 ATAAATCTAATTTGATTTTAAATAATATATATCTATGATGATTTTATTTGGTATTTA 296
Qy 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154
Db 295 TTGGCTAGTGTATAGAGGTGCATTTATTTCTAGATATGATATTAATAGTATTGAC 236
Qy 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeu 174
Db 235 GAATTTAAAT---ACATTAGAAATATTTTAAATAATAAATAAAGTTTATGGAATTA 179
Qy 175 -----AlaProAspPheLysAsn 180
Db 178 ATAGATATGTATAAAAAAATAGTAATAGTAATATGATGTAGTAGTAAATAAATAAAT 119
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu 197
Db 118 GATTATGATAGTATATTTGATAATGAGAAGAAATAGTTTGAATCTTAATTTAATATCTT 59
Qy 198 GluPhe-----LeuGluLysThrTyrLysIleLysLeuTyr 209
Db 58 GAATTTTAAATATCAATTAATCATACAAATTTTGCTACTATAT 17

RESULT 80
US-08-544-332-1/c
; Sequence 1, Application US/08544332
; Patent No. 5935777
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System

; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gerard H. Bencen
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,332
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,867
; FILING DATE: 07-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/107,755
; FILING DATE: 19-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H.
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: UF114.C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8457 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (65..1459)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1474..2151
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (2239..2475)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2502..2987
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3080..6091
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (6277..6768)
US-08-544-332-1
Alignment Scores:
Pred. No.: 0.215 Length: 8457
Score: 105.50 Matches: 56
Percent Similarity: 39.7% Conservativity: 37
Best Local Similarity: 23.9% Mismatches: 82
```

```

Query Match: 6.7% Indels: 59
DB: 2 Gaps: 12

US-10-734-719-9 (1-291) x US-08-544-332-1 (1-8457)
Qy 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCys 45
Db 691 GATCATTTTACTTGTTCGTAATA-----GATAAACACCTTT---GATAAAGATAAA 644
Qy 46 LysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
Db 643 AAAGCTGCATAC-----TTTTTAAATAGTAGTGGTTATATACCAAGAACTTATA 596
Qy 66 IleGlnAsnGlnGluTyr-----GluThrGluLeuLeuMetCysSer----- 79
Db 595 AAACAAAATAAAAAATATGTTTATTGAATCTGATATGACTATATAAAAGCCATAAACAC 536
Qy 80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
Db 535 TACAATAGTACTCCTTAATACTAATTATGCTTATTATATACATTTGATGACTATCAGAATAT 476
Qy 88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 475 TTAATGATATATTTAAATGTAATTTACTTTTAACTCTTTTAACTCTTTGAATTAACAATAT 416
Qy 107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
Db 415 GATAGTCCCGATTGCTGATGTTTAAATAATAATATTTTATATATATATATATTTTCAAC 356
Qy 117 -----AlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyVal 134
Db 355 ATAAATCTAAATTTGAATTTAAATAATATATATTTCTATGAGTTTATTTGGTGATTTA 296
Qy 135 TyrMetCysAlaValAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAsp 154
Db 295 TTGGCTAGTAGTTATAGAGGTGCATTATTTTCTAGATATGATATTAATAGTAGTATTGAC 236
Qy 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeu 174
Db 235 GAATTTAAAT---ACATTAGAAATATTTAATAATAAAAAATAAAAGTTTATGGAATTA 179
Qy 175 -----AlaProAspPheLysAsn 180
Db 178 ATAGATATGTATAAAAAAATAGTAATATGAAATTTATGATGTAGTATAAAATAAAAAAT 119
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu 197
Db 118 GATTATGATAGTTATTTGATAATGAGAGAAATAGTTTAGAATCTAATATTTAATATCTT 59
Qy 198 GluPhe-----LeuGluLysThrTyrLysIleLysLeuTyr 209
Db 58 GAATTTTAAATATCAATTAATCATACAAATTTGTACTATAT 17

RESULT 81
US-09-370-861A-1/c
; Sequence 1, Application US/09370861A
; Patent No. 6410221
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 6410221el Entomopoxvirus Expression System
; FILE REFERENCE: UP114.C4.D1
; CURRENT APPLICATION NUMBER: US/09/370.861A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 07/991,867
; PRIOR FILING DATE: 1992-12-07
; PRIOR APPLICATION NUMBER: US 08/107,755
; PRIOR FILING DATE: 1993-08-19
; PRIOR APPLICATION NUMBER: WO 92/14818
; PRIOR FILING DATE: 1992-02-12
; PRIOR APPLICATION NUMBER: US 07/827,685
; PRIOR FILING DATE: 1992-01-30

; PRIOR APPLICATION NUMBER: US 07/657,584
; PRIOR FILING DATE: 1991-02-19
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 8457
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
US-09-370-861A-1
Alignment Scores: 0.215 Length: 8457
Pred. No.: 105.50 Matches: 56
Score: 39.7% Conservative: 37
Percent Similarity: 23.9% Mismatches: 82
Best Local Similarity: 6.7% Indels: 59
Query Match: 3 Gaps: 12
DB:

US-10-734-719-9 (1-291) x US-09-370-861A-1 (1-8457)
Qy 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCys 45
Db 691 GATCATTTTACTTGTTCGTAATA-----GATAAACACCTTT---GATAAAGATAAA 644
Qy 46 LysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
Db 643 AAAGCTGCATAC-----TTTTTAAATAGTAGTGGTTATATACCAAGAACTTATA 596
Qy 66 IleGlnAsnGlnGluTyr-----GluThrGluLeuLeuMetCysSer----- 79
Db 595 AAACAAAATAAAAAATATGTTTATTGAATCTGATATGACTATATAAAAGCCATAAACAC 536
Qy 80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
Db 535 TACAATAGTACTCCTTAATACTAATTATGCTTATTATATACATTTGATGACTATCAGAATAT 476
Qy 88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 475 TTAATGATATATTTAAATGTAATTTACTTTTAACTCTTTTAACTCTTTGAATTAACAATAT 416
Qy 107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
Db 415 GATAGTCCCGATTGCTGATGTTTAAATAATAATATTTTATATATATATATATTTTCAAC 356
Qy 117 -----AlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyVal 134
Db 355 ATAAATCTAAATTTGAATTTAAATAATATATTTCTATGAGTTTATTTGGTGATTTA 296
Qy 135 TyrMetCysAlaValAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAsp 154
Db 295 TTGGCTAGTAGTTATAGAGGTGCATTATTTTCTAGATATGATATTAATAGTAGTATTGAC 236
Qy 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeu 174
Db 235 GAATTTAAAT---ACATTAGAAATATTTAATAATAAAAAATAAAAGTTTATGGAATTA 179
Qy 175 -----AlaProAspPheLysAsn 180
Db 178 ATAGATATGTATAAAAAAATAGTAATATGAAATTTATGATGTAGTATAAAATAAAAAAT 119
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsn-----ThrAspIleLysAlaLeu 197
Db 118 GATTATGATAGTTATTTGATAATGAGAGAAATAGTTTAGAATCTAATATTTAATATCTT 59
Qy 198 GluPhe-----LeuGluLysThrTyrLysIleLysLeuTyr 209
Db 58 GAATTTTAAATATCAATTAATCATACAAATTTGTACTATAT 17

RESULT 82
US-09-662-254B-26/c
; Sequence 26, Application US/09662254B
; Patent No. 693145
; GENERAL INFORMATION:

```

```
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; TITLE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UF-221C1XC1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
US-09-662-254B-26

Alignment Scores:
Pred. No.:      2.71      Length:    50000
Score:          105.50     Matches:   56
Percent Similarity: 39.7%   Mismatches: 37
Best Local Similarity: 23.9% Indels:    59
Query Match:     6.7%     Gaps:      12
DB:              3

US-10-734-719-9 (1-291) x US-09-662-254B-26 (1-50000)
Qy      26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCys 45
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7507 GATCATTTTACTTGTCGGTAATA-----GATAAACACCTTT---GATAAAGATAAA 7460

Qy      46 LysAlaValaPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
       ||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      7459 AAAGCTGCATAC-----TTTTTAATAGTAGTGGTTATATACCAGAATTTATA 7412

Qy      66 IleGlnAsnGlnGluTy-----GluThrGluLeuIleMetCysSer----- 79
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7411 AACCAAAATAAAAAATATATGTTTATTGAATCTGATATAGACTATAAAAGCCATAAACAC 7352

Qy      80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7351 TACAATAGTACTCCTTAATACTAATATTATGCTTATTTATACATTCGTACTATCAGAATAT 7292

Qy      88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProaspAlaHisLeuGlyTyr 106
       ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      7291 TTAAATGATATATTAAAAATGTAAAATTATTACTTTTTTAATACTTTTGAATTTACAATAT 7232

Qy      107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      7231 GATAGTCCCAGTTGTGGTAGTTTAAATAATAATTTTATATTATATAGTTTATTTCAC 7172

Qy      117 -----AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgileThrSerGlyVal 134
       :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      7171 ATAAATCTAAATTTGAATTTAAAAAATATATATTCTATGAGTTTATTGCGTATTTA 7112

Qy      135 TyrMetCysAlaValalalleLeuLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154
       :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      7111 TTGCGTAGTAGTTATAGAGGTGCCATTTATTTCTAGATATGATATATAATAGTATGCAC 7052

Qy      155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeu 174
       :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      7051 GAATTTAAAAAT---ACATTTAGAAATATTTAATAATAAAAAATAAAAGTTTATGGAATTA 6995

Qy      175 -----AlaProAspPheLysAsn 180
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      6994 ATAGATATGTATAAAAAATAAGTAGTAATATGAATCTATGCTAGTAAAAATAAAAAAT 6935

Qy      181 AspArgSerHisTyrIleGlyHisSerLysAsn-----ThrasplieLysAlaLeu 197
       ||| ||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db      6934 GATTATCATAGTTATATGTAATGATAGAGAAGATAGTTTAGAATCTAATATATTTAATCTCT 6875
```

[illegible]



Qy	13	LeuLysGluIleAspTyrSerArgLeuProAsnAspPhe---AspValPheAerGysAsn	31
Dd	54629	CITTAAGGAA--AAATACTGCACAAAATGATGATGAGTTTGATGACATTTTTAAAGATGAAG	54573
Qy	32	GlnPheTyrPhe-----GluAspLysTyrTyrluengly	42
Dd	54572	CCATTTTATATAAGTGGTATTCTCCAAGAATTAGGCAGAAATTAAACTTTTTATTAAAT	54513
Qy	43	----LysLysCys-----	45
Dd	54512	TTAATTTAAAGACTCGCAGAGATGAACAATAAAAAAGTCGCAATCATNAATTTTGAGTAGA	54453
Qy	46	----LysAlaValPheTyrThrProAsnPhe-----Phe	55
Dd	54452	ACGCCAAGGTCTGTGTAGAGCACTACTCATTTTGTATTTAGCCACACTAAAAGAGCCAGTT	54393
Qy	56	PheGluInlTyTrThrleulysthis-----Leu	65
Dd	54392	TTTGACCACATCATTTGCTTATAGCATAAAAAAATTTGCAGACCAGTTCAAACAATTTTA	54333
Qy	66	IleGlnAsnGlnGluTyrcLutThrGluLeuIleMetCysSerAsnTyrAsnGlnAlahis	85
Dd	54332	AGACACTTAGAAGAATATACAATGATGTAA'TT-----AGTAGA	54294
Qy	86	LeuGluasnGluasnPheVallys-----ThrPheTyrAspTy-PheProAspalahis	103
Dd	54293	ATAGAAGAATTTTCAAGATTTAGGAAGATGCAATATTTTAATCATTAATGGAGATTCA	54234
Qy	104	LeuGlyTyzAspPhePhelysGlnLeuLysGluUpheAsnAlatyPhelysPhetisglu	123
Dd	54233	AGCACTGTTGATATTCAAGAAGNACTAAAAAAA-----CACCCAAACTTTTATGAA	54183
Qy	124	IleTyzPheAsnGlnArgileThrsErGlyValTyrmCyseAlavalalealaleu	143
Dd	54182	CTTTTATATAAACCAAAAAAATTT--GATGGTATTTTATCA- - - - -	54147

```
Qy 144 GlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAla 163
Db 54146 -----TCTCCTCCTTATTTAGGCGAGATTGATTATCATGAGCAG---CATGCCTATGCT 54096
Qy 164 -----PheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 54095 TATGAGCTTTTGCATTTCCAAGATTGATGAGTTGGAGATAGGACCTTAAATTTTAAAGGT 54036
Qy 181 Asp-----ArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAla 196
Db 54035 AGTTCAAAAAGACACAAAAGAGTATATGAAGGCATATCCGATGTTTTTAAATCAACATG 53976
Qy 197 LeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu 216
Db 53975 AAAAGATTTTAAATGAAGCGCAAGATT----- 53946
Qy 217 LeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLys 236
Db 53945 -----TTCAVT-----GTAGTTAAATGATATA 53925
Qy 237 AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLys 256
Db 53924 AAGAACCTATATAAAGAGATTTTTCAGAAAAGTGGCTTATTTAGTTAGAGAGTTTAAA 53865
Qy 257 AsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAsp 276
Db 53864 AGACCGTCTTAAATAGACGAAAGAGATAGAACCCATCTATGAAGCATCTTTGAA 53805
Qy 277 Leu 277
Db 53804 TTA 53802
```

## RESULT 85

```
US-09-601-198-36/c
; Sequence 36, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA
; FILE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 36
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-36
```

## Alignment Scores:

Pred. No.:	0.0274	Length:	1500
Score:	104.00	Matches:	69
Percent Similarity:	36.9%	Conservative:	46
Best Local Similarity:	22.1%	Mismatches:	94
Query Match:	6.7%	Indels:	103
DB:	3	Gaps:	15

US-10-734-719-9 (1-291) x US-09-601-198-36 (1-1500)

```
Qy 27 ValPheArgCysAsnGlnPheTyr-----PheGluAspLysTyr--- 39
Db 1063 GTTTTAAATGACATATTTTATGATAATAAACCAAAATTTTATGCAAAATTTCTATTAA 1004
Qy 40 ---TyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsn-PhePhePheGluGl 58
```

```
Db 1003 GTATATTAGTTTAAAGAGATCCAGAACATTTATTTTCATGAAGCAACAAGATAATTTAAAAA 944
Qy 58 nTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCy 78
Db 943 ATATTATCATCTCTAATAAATTAAGTTCTAAT----- 913
Qy 78 sSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTy 98
Db 912 -----AATAAGATTATTTTATTTTCTTGTGATACAAATTT 881
Qy 98 rPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTy 118
Db 880 CATTTATTTTGTCTATATA-----AATCATTA 854
Qy 118 rPheLysPheHisGluIle-----TyrPheAsnGlnArgIleThrSerGlyValTyrMe 136
Db 853 TTACAAATGCTAGAGAAATATATCAAGTATTACAAATCAAAATTTAATTTAGATCTTCCAG 794
Qy 136 tCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTy 156
Db 793 TAAATATATTAAATCATAGCTATCCAAATAGTCGTATTTATAAAATGATAGGTATTTT 734
Qy 156 rGlnAsnGlySerSerTyrAlaPhe-----AspThrLysGlnGluAsnLe 171
Db 733 CTCTAACGTTAATCGTATATTTTAAAAATAAAACATATGATAAAACAAGATCAACAC 674
Qy 171 uLeuLysLeuAlaProAspPheLys-----AsnAspArgSerHisTyrI 186
Db 673 TTTATTTTTCGCG-----TTTAAAAAACTTAATCAAGCAATGAT----- 634
Qy 186 eGlyHisSerLysAsnThrAspIleLysAlaLeuGlu-----PheLeuGluLysTh 203
Db 633 ----GAAATATAAAATTTAATTTATAGTATTATAAAACAAGAAATTTAATCGATCAAAA 578
Qy 203 rTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLe 223
Db 577 TTATCAATA-----ACTAATAAATTAGATTT 551
Qy 223 uAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspI 243
Db 550 TAAATATCTTTTGAAGAAATCATTTTTTAAATAACAAGAAATATTCAGATCAAAAAATGA 491
Qy 243 eLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysI 263
Db 490 ATTA-----AACGATTTTGAACATTTAATGATCAACTAA 452
Qy 263 eLysIleLys-----GluAsnValTyrTyr----- 271
Db 451 TAAGTTCAATTTATTTGTAATGCATATTGATATGATAATAAATATTATTCGATTAGTAC 392
Qy 272 -----LysLeuIleLysAspLeuLeuArgLeu----- 280
Db 391 ATCTAATACTAAATTAATGAATGATAAAATTCGTTTAGTATTATTAAACATGCTTTGACG 332
Qy 281 -ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 331 AGATTATGATATTACTGATTACATAAAGCTAAA 298
```

## RESULT 86

```
US-09-601-198-108/c
; Sequence 108, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA
; FILE REFERENCE: UAB-13452/22
```

QY	230	AsnPheIlelleGlnGluLysAsnAsn-----TyrThrLysAsp---	242
Db	1740	TACTTATTTTATCAAAATAATAAAACAAACGAAGGTTTTTATGATTTTCATAACTTAAAAAAGCT	1681
QY	243	-----IleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	1680	GATTTTATCTTCTGTAATAATAATGAATC-----AAAAATTACATAAA	1636
QY	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	1635	ACTACATTAAAGATAATCCAGATTTTCATCTTTAAAAAAGATAAAGAT-----	1588
QY	281	ProSerAspIleLysHisTyrPheLys	289
Db	1587	-----AATTATTTTCAA	1576
RESULT 87			
US-09-790-988-1			
; Sequence 1, Application US/09790988			
; Patent No. 6632935			
; GENERAL INFORMATION:			
; APPLICANT: SHIGENOBU, SHUJI			
; APPLICANT: WATANABE, HIDEKI			
; APPLICANT: WATTORI, MASAHIRA			
; APPLICANT: SAKAKI, YOSHIYUKI			
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS			
; FILE REFERENCE: 081356/0159			
; CURRENT APPLICATION NUMBER: US/09/790,988			
; CURRENT FILING DATE: 2001-02-23			
; PRIOR APPLICATION NUMBER: JP2000-107160			
; PRIOR FILING DATE: 2000-04-07			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 640681			
; TYPE: DNA			
; ORGANISM: Buchnera sp.			
US-09-790-988-1			
Alignment Scores:			
Pred. No.:	176	Length:	640681
Score:	103.50	Matches:	72
Percent Similarity:	38.4%	Conservative:	42
Best Local Similarity:	24.2%	Mismatches:	102
Query Match:	6.6%	Indels:	82
DB:	3	Gaps:	17
US-10-734-719-9 (1-291) x US-09-790-988-1 (1-640681)			
QY	2	LysLysValIlelleAlaGlyAsnGlyProSerSerLeuLysGluIleAspTyrSerArgLeu	21
Db	524693	GAAAAATACATACGACCACTGCTCAATGGTGCAAAAAAT	524728
QY	22	ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPhe-----GluAsp	37
Db	524729	-----AGTTTTTAATCTCTTCRAA-----AAATGTATTTTATTGAAAGACAGAAAACAA	524776
QY	38	LysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGlu	57
Db	524777	AAAAAATACTAGCAAAAAATTTTTTAAATTTAGTCATAATGAAATTTTACAAAAGAA	524836
QY	58	GlnTyr-----TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIle	76
Db	524837	ACTTATATTAATTGTTTTATCTCAATTAATTAATAAT-----GTT	524875
QY	77	MetCysSerAsnTyr-----AsnGlnAlaHisLeuGlu-----AsnGluAsnPheValLys	93
Db	524876	CTTTTAGAACAAATATGCAAAAAATATGAATTAATCTCGAAGTCATGATAACGATAAAA	524935
QY	94	ThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLys	113
Db	524936	AAAAATATATCAATTCCTCT-----ATTTTCAAAAAAACCAAT	524974



```
Qy 114 GluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGly 133
Db 524975 AAATTCAGT-----AAAGCGCATATTTAAAT----- 525001
Qy 134 ValTyrMetCysAlaVal-----AlaIleAlaLeuGlyTyrLysGlu 147
Db 525002 ---TATCTTACATCTATATAATCAACTAATCATGAATATATAATAATTAATAAATAA 525058
Qy 148 IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLys 167
Db 525059 ATAAACACAGAAAAATTTAATACATACCATTTCTAAAGCTAATTTTATTTTAAATAAAGAA 525118
Qy 168 GlnGluAsnLeuLysLeuAlaProAsp-----PheLys 179
Db 525119 GAAATAATATATATAAATACTATATCTCAAAAAGAAATAAATAAAGCAATTTGTTAAA 525178
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThr-AspIleLysAlaLeuGluPh 199
Db 525179 ATAGATCCTTCAATCTAT-----AAAAAATAATAACAAATCAAGAGCACAAATT 525229
Qy 199 eLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAs 219
Db 525230 TATTTCAAAA-----AAAATCAAGATAATTTTATATATT-CCAGAAAAAATTTAAATAATA 525282
Qy 219 nPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTy 239
Db 525283 TTTTGTGGAATTAATA-----ACTGATAACTTTTAAATACATTCGCAAAAT----- 525328
Qy 239 rThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAs 259
Db 525329 -----AAAGAGATTATGATTCGTATATAAGAAATATTAC 525363
Qy 259 nPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLys 275
Db 525364 A-----CAATACTCAACTAAAGAAAAAGAGATATAGTATTATTCAA 525406

RESULT 88
US-09-662-254B-47
; Sequence 47, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; APPLICANT: Bawden, Alison Louise
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; TITLE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UF-221C1XC1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 3450
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(3450)
; OTHER INFORMATION:
US-09-662-254B-47
```

```
Alignment Scores:
Pred. No.: 0.118 Length: 3450
Score: 103.00 Matches: 79
Percent Similarity: 35.8% Conservative: 36
Best Local Similarity: 24.6% Mismatches: 94
Query Match: 6.6% Indels: 112
Db: 3 Gaps: 21
```

```
US-10-734-719-9 (1-291) x US-09-662-254B-47 (1-3450)
Qy 5 IleIleAlaGlyAsnGlyProSerLeuLysGluIle-----AspTyrSer 19
Db 451 TTATTAAAGTCAATCAGACCCCTTCATCTAAGTATATATATGGCAATAGTAGATACTATT 510
Qy 20 ArgLeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 39
Db 511 ATGATTTGATAAT-----GAAGATAAAATTT 534
Qy 40 TyrLeuGlyLysLysCysLysAlaValPheTyrThrPro-----AspPhe 54
Db 535 -----GATTTTATACCAATATATTCATTTTATTAATAT 570
Qy 55 PhePheGluGlnTyrTyr-----ThrLeuLysHisLeuIle-----GlnAsn 68
Db 571 TTTTAAATAATATATATATAACATCACAAAAAATATATATTTTAGATCCTAGTAAT 630
Qy 69 GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsn 88
Db 631 CTGGA-----GAAATAAAGATGTACTAAGCTAATTTAATCGACGAAAGTTTAAAAAT 687
Qy 89 Glu-----AsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
Db 688 AAATATAATAATTTTACAAAGAAATTA---TCATATTTTGATATATCAATAGTAGATAT 744
Qy 107 AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPhe 126
Db 745 AATTTCTATAAATGATGGTGGTGAATTTTAAATAATTTTAA-----GATATCAATACT 795
Qy 127 AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys 146
Db 796 AATAAAAAATATTATTGAAAAATTA-----GATGTAATTTAATAATAATATTATAAAA 846
Qy 147 GluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThr 166
Db 847 TCAATATATCTATATAACATA-----ATGATACA 876
Qy 167 LysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArg----- 182
Db 877 AATGTAGAAGATATATAATAATA-----ATAATGAACGATACAAATTTATTATTA 927
Qy 183 -----SerHisTyrIleGlyHisSer----- 189
Db 928 TTGAATGAAATATATAGTGAATTTTACCAAACTCAAGCAAAATATATATTTAGTGGGA 987
Qy 190 -----LysAsnThrAspIleLysAlaLeuGluPhe 199
Db 988 TTACGTCGCATTATATATGAAAAAGCAACAAATAAAAAATATTAGCAATTTATATATG 1047
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 1048 TTAGATTTCATTTGTAAGTATATATTATAT-----TTATTAGAAAGA 1089
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys----- 236
Db 1090 TATTACGAA-----AATGATATAACCACTTAACTTAATGAATCTAAAGATTA 1134
Qy 237 ---AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255
Db 1135 ATAAAAACAATATTATTAAGATAATTTA-----AATTC 1167
Qy 256 LysAsn---IleAsnPheLysLysIleLys-----IleLysGluAsnValTyrTyrLysLeu 273
Db 1168 AAAAATAGCGTTAAATTTGGATTCATATAAATATTTTAAAGAAAAATCAATAATAATATT 1227
Qy 274 Ile 274
Db 1228 ATT 1230
RESULT 89
US-09-601-198-70
```

```
; Sequence 70, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 70
; LENGTH: 2427
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-70

Alignment Scores:
Pred. No.: 0.0815 Length: 2427
Score: 102.50 Matches: 74
Percent Similarity: 38.1% Conservatives: 57
Best Local Similarity: 21.5% Mismatches: 123
Query Match: 6.6% Indels: 90
DB: 3 Gaps: 16

US-10-734-719-9 (1-291) x US-09-601-198-70 (1-2427)

Qy 3 LysVallelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeuPro 22
Db 940 AAATACCTACTATTGTTGTAATATTATTAATAAATGAAGTATTGATAAAACCGTTATATA 999

Qy 23 AsnAspPheAspValPheArgCysAsnGlnPheTyr-----PheGluAspLysTyr 39
Db 1000 AATGAATTAATAAATCGATACTATAAACAAGAGTATTAGATTATTATTCAAAACCAAGTT 1059

Qy 40 TyrLeuGlyLys----- 43
Db 1060 TTAATTCAAAAGAAATTAATTGACGATCCAAAAAATAAGAGAGTGCACAAAAATATT 1119

Qy 44 ---LysCysLysAlaValPheTyrThr-----ProAsnPhePheGlu 57
Db 1120 GCAAAAGCCAAAGCCTTGTAAATGCTTAAATTCGTGAAGAAACAAATATTATTTATGAA 1179

Qy 58 -----GlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeu 75
Db 1180 AACTCACAAAATCAGCTCTTGAATAAATAAATAAGATTCGTGATTATGAGATCAATTA 1239

Qy 76 -----IleMetCysSerAsnTyrAsnGln-----AlaHisLeu 86
Db 1240 AAACAAATAAATAATTGCTCTACAAAATTCGAATATTCTTAATAACGAACCTTATTCAAATTA 1299

Qy 87 GluAsnGluAsnPheVal-----LysThrPheTyrAspTyrPhe 99
Db 1300 AAAAATGAACAAAAGTATTAGCAGCAATCTCGATCAAGCAAAATTTAAACACTTTTATA 1359

Qy 100 ProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 1360 TATATATGTTAATTAACACAGATTTCCGTGCGATTATTAATAATAATTAATTAATTTATAT 1419

Qy 120 LysPheHisGlu-----IleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCys 137
Db 1420 CGCAACATGAATCAGCAATTTAATAATGATGTAAAGAGTATTCAAATTTTATTTGTCACA 1479

Qy 138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln 157
Db 1480 AATGCA-----AGCAATCCACATATATACAAA 1506

Qy 158 AsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProasp 177
Db 1507 AACGGTTTAATTGACCAAGAAATCAATGAATTGATATATATGTCGCAATTAATGAAAAA 1566

Qy 178 PheLysAsnAspArgSerHis-----TyrIleGlyHisSerLys 190
Db 1567 AATAAAAATAGTCAAAATCCAAATAAAAAATTAAGCGCAATTTTATTATTAATGATGATAAT 1626

Qy 191 AsnThrAspLysAlaLeuGluPheLeuGluLysThrTyrLysIleLeuLysLeuTyrCys 210
Db 1627 TCAAAATAAAATTAATAAACTTTTATGACCTTTTATCAAAATTAC---ATTAAAGTCATT--- 1680

Qy 211 LeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn 230
Db 1681 -----CAAATTTTAAAAAATAGTTATGAAAAAGATCCAGATGATAAAGAT 1728

Qy 231 PheIleIleGlnGluLysAsnAsnTyrThrLys----- 241
Db 1729 TTACTAGTTGAAGACGT---GATTATATCTCTTGAATTAATACATCACTAATTAATATA 1785

Qy 242 -----AspIleLeuIleProSerSerGluAlaTyrGlyLysSh 254
Db 1786 TATCAACTTCACCAATTTTAAATATCTTATTTAATCAATCTGAATGATTATATCAAT 1845

Qy 254 eSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIle 274
Db 1846 C-----AATATCTAAAGAAATGAGTGATTATAAAA---AATATTTACTACAAATTTT 1896

Qy 274 eLysAspLeu 277
Db 1897 GAACCATATC 1906

RESULT 90
US-09-830-230A-348
; Sequence 348, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 348
; LENGTH: 1779
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-830-230A-348

Alignment Scores:
Pred. No.: 0.0784 Length: 1779
Score: 101.00 Matches: 48
Percent Similarity: 39.9% Conservatives: 35
Best Local Similarity: 23.1% Mismatches: 80
Query Match: 6.5% Indels: 45
DB: 3 Gaps: 8

US-10-734-719-9 (1-291) x US-09-830-230A-348 (1-1779)

Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAsp 97
```

Db 150 TGTCTCTAGTCGAATTAAGTCTTATTTATGAGGAGTCTTATTTCTCAAGTCTTATTTTGTGAC 209  
Qy 98 Tyr-PheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAl 117  
Db 210 TACAAGTAAACATCTTTAAGTCTTATTTATTTATTTTCAAAAATTTTGAATAATAGTAGG 269  
Qy 117 atyRPhelysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCy 137  
Db 270 AATTGATAAATTTAAT-----ACCGGTGCATATATTAC 302  
Qy 137 sAlaValAlaIleAlaLeuGly---TyrLysGluIleTyrLeuSer-----G1 152  
Db 303 ATCTAGTCCCTTTCTCAAGGAGATTACAAGCGTATTGCTATTTGGAACCTGCGATTTCATGG 362  
Qy 152 yIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnLeuAsnLeuLe 172  
Db 363 TATTATCTTAGTGTAAATGAGCT-----ATTAGTCTTAAAAATTTTAAA 407  
Qy 172 uLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnTh 192  
Db 408 TCGTTTGATTCGCGAGATT-----TATTAGGTGCAGGATATTACGA 449  
Qy 192 rAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCy 212  
Db 450 TATTATTAGTGTATTGAATTTTCAAAAAGAG-----483  
Qy 212 sProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn---Ph 231  
Db 484 -----ACAAATAATTTATATTTTCTCTGAGTATTGAGAGATATTTT 527  
Qy 231 eIleIleGlnLysAsnAsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTy 251  
Db 528 TTTAATTAGTCAGAAAAGTGGATTATTAAAAAATATCTTTTCTTCAAAAAGCAAAAT 587  
Qy 251 rGlyLysPhe-----SerLysAsnIleAsnPheLysIleLysIleLysG1 267  
Db 588 AATAGTATTTTAGACTTATCTAGTAAGAACTAGAAAAATTTTAGTCAGAACATATGA 647  
Qy 267 uAsnValTyrTyrLysLeuLe 274  
Db 648 CAATCAATTTTATCTTATATT 669

## RESULT 91

US-09-830-230A-347  
; Sequence 347, Application US/09830230A  
; Patent No. 6902893  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Lyme Disease Vaccines  
; FILE REFERENCE: PB481US  
; CURRENT APPLICATION NUMBER: US/09/830,230A  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: PCT/US98/12718  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/057,483  
; PRIOR FILING DATE: 1997-09-03  
; PRIOR APPLICATION NUMBER: 60/053,344  
; PRIOR FILING DATE: 1997-07-22  
; PRIOR APPLICATION NUMBER: 60/053,377  
; PRIOR FILING DATE: 1997-07-22  
; PRIOR APPLICATION NUMBER: 60/050,359  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 756  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 347  
; LENGTH: 1839  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-830-230A-347

Alignment Scores:

Pred. No.:	0.0822	Length:	1839
Score:	101.00	Matches:	48

Percent Similarity: 39.9%  
Best Local Similarity: 23.1%  
Query Match: 6.5%  
DB: 3  
Conservative: 35  
Mismatches: 80  
Indels: 45  
Gaps: 8

US-10-734-719-9 (1-291) x US-09-830-230A-347 (1-1839)

Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAsp 97  
Db 210 TGTCTCTAGTCGAATTAAGTCTTATTTATGAGGATGTTTCTTCAAGTCTTATTTTGTGAC 269  
Qy 98 Tyr-PheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAl 117  
Db 270 TACAAGTAAACATCTTTAAGTCTTATTTATTTTCAAAAATTTTGAATAATAGTAGG 329  
Qy 117 atyRPhelysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCy 137  
Db 330 AATTGATAAATTTAAT-----ACCGGTGCATATATTAC 362  
Qy 137 sAlaValAlaIleAlaLeuGly---TyrLysGluIleTyrLeuSer-----G1 152  
Db 363 ATCTAGTCCCTTTCTCAAGGAGATTACAAGCGTATTGCTATTTGGAACCTGCGATTTCATGG 422  
Qy 152 yIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLe 172  
Db 423 TATTATCTTAGTGTAAATGAGCT-----ATTAGTCTTAAAAATTTTAAA 467  
Qy 172 uLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnTh 192  
Db 468 TCGTTTGATTCGCGAGATT-----TATTAGGTGCAGGATATTACGA 509  
Qy 192 rAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCy 212  
Db 510 TATTATCTTAGTGTATTGAATTTTCAAAAAGAG-----543  
Qy 212 sProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn---Ph 231  
Db 544 -----ACAAATAATTTATATTTTCTCTGAGTATTGAGAGATATTTT 587  
Qy 231 eIleIleGlnLysAsnAsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTy 251  
Db 588 TTTAATTAGTCAGAAAAGTGGATTATTAAAAAATATCTTTTCTTCAAAAAGCAAAAT 647  
Qy 251 rGlyLysPhe-----SerLysAsnIleAsnPheLysIleLysIleLysG1 267  
Db 648 AATAGTATTTTAGACTTATCTAGTAAGAACTAGAAAAATTTTAGTCAGAACATATGA 707  
Qy 267 uAsnValTyrTyrLysLeuLe 274  
Db 708 CAATCAATTTTATCTTATATT 729

## RESULT 92

US-09-662-254B-7  
; Sequence 7, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bowden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1XCI  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 3318  
; TYPE: DNA



Db 1078 AGT---TTCACCTTCAAACTTATAATATTTCTGAA----- 1110  
Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheVal-----LysThrPhe 95  
Db 1111 -----TCTCGCGCTACAGTTTCTAATAATCTTTTCATTGTAATGAATGCCCAT 1158  
Qy 96 TyrAsp-TyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPh 115  
Db 1159 TACATCTATTTTCAAAATGTTCTATTG-----GAGTTTGTAGAAATTTGAAC----- 1207  
Qy 115 eAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyVal-- 134  
Db 1208 -----GTTAAACAGCAAGTTGTTGCTGGAATGAT 1236  
Qy 135 -----TyrMetCysAlaValAlaLeuLeuGlyTyrLysGluLeuTyrLeuSerGlyI 153  
Db 1237 GTACTATATAACACTTGGCGCAATGATGCTGGAAAGAAAGAAATATATGAACCTTAAGAT 1296  
Qy 153 e-----AspPheTyrGlnAsnGlySerSe 161  
Db 1297 TTGGGTGAGGAATGGGAGGACTTCAAGAAAGTTGTAGAATTCAGCTTGTGTGATGA 1356  
Qy 161 rTyrAlaPheAspThrLysGlnGluAsnLeuLysLeu-----AlaProAs 177  
Db 1357 TAGTGCA-----AAACTGGGGCATTTATCAATGTTCCAAACCCAAACAGCCCGA 1407  
Qy 177 pPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLe 197  
Db 1408 GTTCCAGATCTTGCTCGCTTTTGTCTGTTTCAGGATTTATAATAACACAGGTCAATATAT 1467  
Qy 197 u-----GluPheLeuGluLysThrTyrLysI 206  
Db 1468 ATAATGACTTACTTTTACTTTCTTTTCTTTTGTGTTTAAATTTCCATTAACACCTATAAT 1527  
Qy 206 eLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAs 226  
Db 1528 TCAATCTTTAATCTATTGCGAGTACTAATAATCTTTCAATTT----- 1570  
Qy 226 nLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIlePr 246  
Db 1571 -----GAACAAATGTCATATACA-----ATCTA 1593  
Qy 246 oSerSerGlu-AlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleL 266  
Db 1594 CTTTTCAGAAATGCTCAATTTGGAGTTTGTAGAAATTTGAATGTGAAGAACAACTTGTT 1653  
Qy 266 ysGluAsnValTyrTyr 271  
Db 1654 CTGGAATGATGACTAT 1670

RESULT 94  
US-09-531-120-211/c  
; Sequence 211, Application US/09531120  
; Patent No. 6972197  
; GENERAL INFORMATION:  
; APPLICANT: PREUSS, DAPHNE  
; APPLICANT: COPENHAVEN, GREGORY  
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS  
; FILE REFERENCE: ARCD-309US  
; CURRENT APPLICATION NUMBER: US/09/531,120  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/125,219  
; PRIOR FILING DATE: 1999-03-18  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 211  
; LENGTH: 1082144  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-531-120-211

Alignment Scores:  
Pred. No.: 834 Length: 1082144

Score: 100.50 Matches: 81  
Percent Similarity: 33.4% Conservative: 35  
Best Local Similarity: 23.3% Mismatches: 106  
Query Match: 6.4% Indels: 126  
DB: 4 Gaps: 20  
US-10-734-719-9 (1-291) x US-09-531-120-211 (1-1082144)  
Qy 15 GluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPheTyr 34  
Db 463809 GAAATATATATGGACGT---CCATATCTCTGTGACCAATTAATAAAGAAATTTCT 463753  
Qy 35 PheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhe 54  
Db 463752 TTC-----TTTTTT-----GTTTTTTTGACATCAAAATTTA 463723  
Qy 55 ---PhePheGluGln-----TyrTyrThrLeuLysHisLeu----- 65  
Db 463722 TCATTTCTTTTCTCAGAACTTAAAGTCTCTTATATAACAACCTTAAAGGAGCTTAAGATGATC 463663  
Qy 56 -----GlnGlnAsnGln 69  
Db 463662 TCCCGCCTAAAGTCAAGACAGCGTTGCTTTTGTCTGGATATAGTAGTTAGCACCAAT 463603  
Qy 70 GluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGlu 89  
Db 463602 CGTCTTTCTCTTACGTGGATTGCTGCGCAACACTTCAGCTTTAGCCACCTCTCTGTACAAC 463543  
Qy 90 AsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeu----- 104  
Db 463542 CACTTCTGCCACTTTGCTTACAGCACCTCTGCTACAGCACCTTTTGCTTCAACGCTTTC 463483  
Qy 105 -----GlyTyrAsp----- 107  
Db 463482 CTCTTCAACAATGTCCTGCGGTTTCGATTTCAGGCTTCGATTAGTTTCAAGAGGATGCA 463423  
Qy 108 -----PhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119  
Db 463422 GTCAATTCGTATACTAGTATCATCATGTTTCCAAAGATCAAGAGTTTTCGGTATATCAT 463363  
Qy 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal 139  
Db 463362 TGGTTC-----TTGTTTATGTTTGGCAGGATGATCAGGTTCTACTAGCTGCAGG--- 463312  
Qy 140 AlaIleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPhe----- 155  
Db 463311 -----GGATTTCAGAAAG-----TGCAAGGCAATGATATAAATAATTCCAAT 463270  
Qy 156 -----TyrGlnAsnGlySerSerTyr-AlaPheAspThrLysG 168  
Db 463269 AGATGATAAAGATGTACAACTATCATCGAAGAAAGTGGTCTGCATTTTCTCAA----- 463215  
Qy 168 nGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHi 188  
Db 463214 -----GCACCCCAACAGAAAGAAACCTTAA----- 463191  
Qy 188 sSerLysAsnThrAspIleLysAlaLeuGluPheLeuGlyLysThrTyrLysIleLysLe 208  
Db 463190 -----AAGAATATGAGTGTTCATCAACAAACTAC---CGTAAACA---TCTATAAGTTA 463141  
Qy 208 uTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIle-----GluLeuAl 224  
Db 463140 ATATTGTTTGGACCGGAGAAATTAATTAATTATGATTAATAATAATTAATTAATTAATTA 463081  
Qy 224 aProAsnLeuAsnAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLe 244  
Db 463080 TAAAGAGATTTT-ACAAATTTTATAATTTCTTACAAAAAATTTAATTTTAAAGTGAICTT 463022  
Qy 244 u-----ileProSerSerGluAlaTyrGlyLysPh 254  
Db 463021 TTCTTTTATAATCAATAATCTTTCTTAAATCCGATTATAAGTAGAAGAAACAATTTAATTAAGTT 462962  
Qy 254 eSerLysAsnIleAsnPhe-----LysLysIleLysIleLysGluAsnVa 269





ZIP: 32606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/544,332  
 FILING DATE: 19-AUG-1993

CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/991,867  
 FILING DATE: 07-DEC-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/107,755  
 FILING DATE: 19-AUG-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: WO 92/14818  
 FILING DATE: 12-FEB-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/827,685  
 FILING DATE: 30-JAN-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/657,584  
 FILING DATE: 19-FEB-1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Bencen, Gerard H.

REGISTRATION NUMBER: 35,746  
 REFERENCE/DOCKET NUMBER: UF114.C4

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 904-375-8100  
 TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 1511 base pairs  
 TYPE: nucleic acid

STRANDEDNESS: double  
 TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)  
 ORIGINAL SOURCE:

ORGANISM: Ansacta moorei entemopoxvirus

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: complement (18..218)

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: complement (234..782)

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 852..1511

US-08-544-332-8

Alignment Scores:  
 Pred. No.: 0.0813 Length: 1511

Score: 100.00 Matches: 74

Percent Similarity: 36.3% Conservative: 39

Best Local Similarity: 23.8% Mismatches: 100

Query Match: 6.4% Indels: 99

DB: 2 Gaps: 14

US-10-734-719-9 (1-291) x US-08-544-332-8 (1-1511)

Qy 1 MetLysLysValIleAlaGlyAsn--GlyProSerLeuLysGluIleAspTyrSer 19

Db 727 ATCAATTCTGTGTTGTTTCCCGAAGAACATAGACCAATTAT----- 768

Qy 20 ArgLeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPhe----- 35

Db 769 -----AATTCTATCGACATTTTATTATTATTGATATATTTTTCACAAAAA 819

Qy 36 -----GluAspLysTyrTyrLeuGlyLysCysLysAlaValPhe-TyrThrPr 52

Db 820 TTAATCAATGAAAAAATAAATAATTCATAAATGGATTTTACTAAATTCGATATAATTT 879

Qy 52 oAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrG1 72  
 Db 880 TAATAAATATTTA--AAATATTATTAATTTAAAAAATAATAATAACAGAGAT----- 932  
 Qy 72 uThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheVa 92  
 Db 933 -----AATGTTTATTAATAATAATATT 955  
 Qy 92 lLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLe 112  
 Db 956 AAAAAAATTAGTTAAT-----TT 973  
 Qy 112 uLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSe 132  
 Db 974 AGAAGAA-----TTGCATATATATATATATATATATATATATATTTTAAA 1015  
 Qy 132 rGlyValTyrMetCysAlaValAlaLeuGlyTyrLysGluIleTyrLysGluIleTyrLeuSerG1 152  
 Db 1016 TAAT-----ATTCCAGAAAAATATAAAGTTTATATATTTTCAAA 1054  
 Qy 152 yLleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLe 172  
 Db 1055 TTTA-----AATATTATTAATTTTATAACAAAAATTTAAAAATATAAAC 1102  
 Qy 172 uLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnTh 192  
 Db 1103 ATATTTAGAT-----ATACTTATATAACAAAAATAG 1132  
 Qy 192 rAspIleLys-----AlaLeuGluPheLeuGluLysThrTyrLysI1 206  
 Db 1133 CAATATAAGTAATATATATACTACCATCTCTAGAAATTTTAAAT----- 1178  
 Qy 206 eLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAs 226  
 Db 1179 -----TGTGAATCATGTAATATAATAAATGACTATAATTTTATTAATAATTTTAGTAAA 1228  
 Qy 226 nLeuAsnSerAsnPheIleIleGlnGluLys-----AsnAsnTyrThrly 241  
 Db 1229 TTTAAAAAATAATAATATCTAAAAATAAATTTGGTAACTTTAATAATGTTTTCCTAT 1288  
 Qy 241 sAspIleLeuIleProSerSerGluAla-----TyrGlyLysPheSerly 256  
 Db 1289 TAGTATAGTTCAGTTAAATATGGAATCAATACAAATAAAGATTATAAATTTATAGAAA 1348  
 Qy 256 sAsnIleAsnPheLysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAs 276  
 Db 1349 ATTAATTAATTTAAAAAATAATAGATATATCTTTCAATGTTAAAAAATAATATA---CA 1405  
 Qy 276 pLeuLeuArgLeuProSerAspIleLysHis 286  
 Db 1406 TTTGATAAAATTTCCAAAAAGTATAAATCAAT 1436  
 RESULT 98  
 US-09-370-861A-8  
 ; Sequence 8, Application US/09370861A  
 ; Patent No. 6410221  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moyer, Richard W.  
 ; APPLICANT: Hall, Richard L.  
 ; APPLICANT: Gruidl, Michael E.  
 ; TITLE OF INVENTION: No. 6410221el Entomopoxvirus Expression System  
 ; FILE REFERENCE: UF114.C4.D1  
 ; CURRENT APPLICATION NUMBER: US/09/370,861A  
 ; CURRENT FILING DATE: 1999-08-09  
 ; PRIOR FILING DATE: 1999-08-09  
 ; PRIOR FILING DATE: 1992-12-07  
 ; PRIOR APPLICATION NUMBER: US 08/107,755  
 ; PRIOR FILING DATE: 1993-08-19  
 ; PRIOR APPLICATION NUMBER: WO 92/14818  
 ; PRIOR FILING DATE: 1992-02-12  
 ; PRIOR APPLICATION NUMBER: US 07/827,685  
 ; PRIOR FILING DATE: 1992-01-30



; PRIOR APPLICATION NUMBER: US 07/657,584

; PRIOR FILING DATE: 1991-02-19

NUMBER OF SEO ID NOS: 78

; SOFTWARE: PatentIn version 3.1

; SEO ID NO 8

```

; SEQ ID NO 8
: LENGTH: 1511

```

```

; LENGTH: 13
; TYPE: DNA

```

ORGANISM: Amsacta moorei entomopoxvirus

; ORGANISM: AMIN  
IIS-09-370-861A-8

Alignment Scores:

Alignment Scores:			
Pred. No.:	0.0913	Length:	1511
Score:	100.00	Matches:	74
Percent Similarity:	36.3%	Conservative:	39
Best Local Similarity:	23.8%	Mismatches:	100
Query Match:	6.4%	Indels:	99
DB:	3	Gaps:	14

US-10-734-719-9 (1-291) x US-09-370-861A-8 (1-1511)

Qy	1	MetLysLysValIleIleAlaGlyAsn---GlyProSerLeuLysGluLeuAspTyrSer	19
Db	727	ATCAATTCTGTTGTTTGGCAGAAAAATAGACCAATTATT---	768
Qy	20	ArgLeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPhe---	35
Db	769	-----AATCTCATCGACATTTTTTTTTTATTATTTCGATATATTTTTTCAAAAAAAA	819
Qy	36	-----GluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPhe---TyrThrPr	52
Db	820	TTAATCAATGAAAAAAAATAAAATATCAAAATGGATTTACTAAATTTCTGAATAATTT	879
Qy	52	oAsnPhePhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGlnTyrGl	72
Db	880	TAATAAATATTTTA--AAATATTATAATTTAAAAAAAATAATAATAAACAGAGAT	932
Qy	72	uThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluGluAsnGluAsn	92
Db	933	-----AATGTTATTAAATATTATAATATT	955
Qy	92	lLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGln	112
Db	956	AAAAAATTTAGTTAAAT	973
Qy	112	uLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSe	132
Db	974	AGAAGAA-----TTGCATATAATATATATGATAATAATATTTTAA	101
Qy	132	rGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGl	152
Db	1016	TAAT-----ATCCAGAAATATTTAAAGTTTATATATTTTCA	105
Qy	152	yIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeu	172
Db	1055	TTTA-----AATATTATTAAATTTTATACAAAATTTAAAAAATAATAC	110
Qy	172	uLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsn	192
Db	1103	ATATTTAGAT-----ATATCTTATACAAAAATAG	113
Qy	192	rAspIleLys-----AlaLeuGluPheLeuGluLysThrTyrLysI	206
Db	1133	CAATATAAGTAATATTATACACACATTTCTATAGAAATTTTAAAT	117
Qy	206	eLysLeuTyrCysLeuCysProAsnSerLeuAlaAsnPheIleGluLeuAlaProAs	226
Db	1179	-----TGTGAATCATGTAATAATAAATGACTATAATTTTATTATAATTTTAGTAA	122
Qy	226	nLeuAsnSerAsnPheIleIleGlnLys-----AsnAsnTyrThrLys	241
Db	1229	TTTAAAAAATAATAATATCTAAAAATAAATTTTGGTAACTTTAAATAATTTTCTT	128
Qy	241	sAspIleLeuIleProSerSerGluAla-----TyrGlyLysPheSerLys	256

APPLICATION NUMBER: PCT/US94/07902  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Saliwanchik, David R.  
 REGISTRATION NUMBER: 31,794  
 REFERENCE/DOCKET NUMBER: MA79  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 904-375-8100  
 TELEFAX: 904-372-5800  
 INFORMATION FOR SEQ ID NO: 27:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3543 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 PCT-US94-07902-27

Alignment Scores:  
 Pred. No.: 0.313 Length: 3543  
 Score: 99.50 Matches: 53  
 Percent Similarity: 31.3% Conservative: 24  
 Best Local Similarity: 21.5% Mismatches: 100  
 Query Match: 6.4% Indels: 69  
 DB: Gaps: 9

US-10-734-719-9 (1-291) x PCT-US94-07902-27 (1-3543)

QY	118	TyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCys	137
DB	688	TATTTAAAAA-----AACATCGCAATTCGAT-----TATTAGAG	723
QY	138	AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln	157
DB	724	CCTTTGCCAACGCAATTCGATTTATTCAGTAAAGCTATAGAGATTACACT	783
QY	158	AsnGlySerTyrAlaPheAspThr-----LysGlnGluAsnLeuLeuLysLeuAla	175
DB	784	AAT-----TATTGTGTAACAACCTTATAAAAAAGGATTAAATTTAATTAACGACG	834
QY	176	ProAspPheLysAsnAspArgSer-----	183
DB	835	CCTGATAGTAATCTTGATGGAATAATAAACTGGAACACATACATACGTAATCGAACAAA	894
QY	184	-----HisTyrIleGlyHisSerLysAsnThrAspIle-----	194
DB	895	ATGACTACTGCTGATTAGATCTTGTTCACCTTTCTCCTAAATTTATGATGAGTAAATAT	954
QY	195	-----LysAlaLeuGluPheLeuGlu	201
DB	955	CCAATAGGTGTCGAATCTGACTCGAGAAATTTATCAGGTACTTAACCTCGAAGAA	1014
QY	202	LysThrTyrLysIleLysLeuTyrCysLeuPheProAsnSerLeuLeu-----	217
DB	1015	AGCCCTATATAATATTATGACCTTCAATATCAAGAGGATTCACTTACGTCAGACCGCAT	1074
QY	218	-----AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn	230
DB	1075	TTATTTACTGGCTTGATTTCTTTGAAATTTTATGAAAAAGCGCAAACTACTCTCTAATAAT	1134
QY	231	PheIleIleGlnGluLysAsn-----AsnTyrThrLysAspIleLeuLeuProSerSer	248
DB	1135	TTTTTCCACGACCATTAATATATGTTTCATATCACACTTGTGTAATAATATATCCCAAAATCT	1194
QY	249	GluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsn	268
DB	1195	AGTGTGTTTGGAAATCACAATGTAACGTAAATAATAAATCTCTGTTGGCAACAAT	1254
QY	269	ValTyrTyrLysLeuLeu 274	
DB	1255	ATTATATTTTTTATTA 1272	

RESULT 100  
 PCT-US94-07902-27  
 Sequence 27, Application PC/TUS9407902  
 GENERAL INFORMATION:  
 APPLICANT: Street address: 4980 Carroll Canyon Road  
 APPLICANT: City: San Diego  
 APPLICANT: State/Province: California  
 APPLICANT: Country: US  
 APPLICANT: Postal code/zip: 92121  
 APPLICANT: Phone number: (619) 453-8030  
 APPLICANT: Telex number:  
 TITLE OF INVENTION: Materials and Methods for the Control of  
 TITLE OF INVENTION: Calliphoridae Pests  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: David R. Saliwanchik  
 STREET: 2421 N.W. 41st Street, Suite A-1  
 CITY: Gainesville  
 STATE: FL  
 COUNTRY: USA  
 ZIP: 32606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:

Qy 249 GluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysLysIleLysLysLysGluAsn 268  
Db 1195 AGTGTGTTTGGAAATCACAATGTAATGTAATAATTAAATCTCTGGTTGGCAACAAT 1254  
Qy 269 ValTyrTyrLysLeuLeu 274  
Db 1255 ATTATATATTTTATTTATTA 1272

RESULT 101  
US-09-662-254B-24/c  
; Sequence 24, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Baden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1X1  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-662-254B-24

Alignment Scores:  
Pred. No.: 13.6 Length: 50000  
Score: 99.50 Matches: 61  
Percent Similarity: 36.4% Conservative: 35  
Best Local Similarity: 23.1% Mismatches: 103  
Query Match: 6.4% Indels: 65  
DB: 3 Gaps: 13

US-10-734-719-9 (1-291) x US-09-662-254B-24 (1-50000)

Qy 66 IleGlnAsnGlnGluTyr-----GluThrGluLeuIleMetCysSerAsnTyrAsnGln 83  
Db 26584 ATACAAAATAATATATATATATGTTAAATAATGATCTAAATAATAATTCAAAATTTATAATAAA 26525  
Qy 84 AlaHisLeuGlu----- 87  
Db 26524 TGGAAAAATGAATTAATGATATATTAAATTTTAAATAATAATAAGAAAAATAATA 26465  
Qy 88 -----AsnGluAsnPhe----- 91  
Db 26464 CTTTATTTCATACAAATATATATATATATAGTTTAAATAATAATTTATAATAATGCAA 26405  
Qy 92 ---ValLysThrPheTyrAspTyr-----PheProAspAlaHisLeuGly 105  
Db 26404 GAAATTTAAATAATTTATGATTTTAAACATATATTTTCCCGAATTCATAATAAA 26345  
Qy 106 TyrAspPhePheLysGlnLeu-----LysGluPheAsnAlaTyrPhe 119  
Db 26344 TATAACTATATTTTCACATTTATTTTTCCTAATAATAATAATTTTCAATCATATATT 26285  
Qy 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal 139  
Db 26284 GATTTTGAT-----TATGTAAAAAATAATAATAATTTTAAATAATTTATATACCT 26231  
Qy 140 AlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159  
Db 26230 GTTTACAAACTATATTGAAAAATATGAT-----ATTGTTATAATCAAAATAGT 26180  
Qy 160 SerSerTyrAlaPheAspThrLysGln-----GluAsnLeuLeuLysLeuAlaProAsp 177  
Db 177 ----- 177

Db 26179 AATAAAATATATTTTACATAATAAAGAAATATATAATACAGTATTGAATTAATAATGAT 26120  
Qy 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu 197  
Db 26119 TTTTAAATAGAT-----GGAATAGATATAATAATAATATATATAACTATT 26075  
Qy 198 GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu 217  
Db 26074 AGATCAACGGATCTACAATTACTTATTCTGCATACGCATATGCA-----ACAATATTA 26021  
Qy 218 AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn-----PheIleIleGlnGlu 235  
Db 26020 TATGATTTTACCATATAGATTAGCAAAATTTAGATATTAATCAATATTTGGAATTTAGAA 25961  
Qy 236 LysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255  
Db 25960 AGTTCTAATATATTA-----GGAATATTTATCTACATATGAAGAACAAAAAGTTTCT 25904  
Qy 256 LysAsnIleAsnPhelLysLysIleLysLysGluAsnValTyrTyrLysLeuIleLys 275  
Db 25903 AAATATATTAAT-----AATATAGAAATAGAAAAAATATATATTTAAATTAAGGAA 25850  
Qy 276 AspLeuLeuArg 279  
Db 25849 TCCTAACCTTAGA 25838

RESULT 102  
US-07-991-867B-32  
; Sequence 32, Application US/07991867B  
; Patent No. 5476781  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Hall, Richard L.  
; APPLICANT: Gruidl, Michael E.  
; TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/991,867B  
; FILING DATE: 12-DEC-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/14818  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,685  
; FILING DATE: 30-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/657,584  
; FILING DATE: 19-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UF114.C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 660 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double

RESULT 103  
US-08-107-755A-32  
; Sequence 32, Application US/08107755A  
; Patent No. 5721352  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Hall, Richard L.

Db 208 AATATTATTAAATTTTATAACAAAATTAATAACATATTTAGAT----- 261  
Qy 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLys----- 195  
Db 262 -----ATATCTTATAACAAAATAGCAATATAGTAATATT 297  
Qy 196 -----AlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211  
Db 298 ATACTACCACATCTATAGATTTTAAAT-----TGTGAA 333  
Qy 212 CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231  
Db 334 TCATGTAATAATAAGTACATATAATTTTATTATAATTTAGTAAATTTAAAAAATTAATA 393  
Qy 232 IleIleGlnGluLys-----AsnAsnTyrThrLysAspIleLeuPro 246  
Db 394 ATATCTAAAAATAAATTTGGTAACTTTTAATTAATGTTTTCTTATTAGTATAGTTGAGTTA 453  
Qy 247 SerSerGluAla-----TyrGlyLysPheSerLysAsnIleAsnPheLys 261  
Db 454 AATATGAATCAATACAAATAAAAGATTATAATTTATAGAAAAATTAATTTAAAA 513  
Qy 262 LysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuPro 281  
Db 514 AAATTAGATATATCTTCAATGTTTAAAAAATAATATA---CATTTGATAAAATTTCCA 570  
Qy 282 SerAspIleLysHis 286  
Db 571 AAAAGTATAACTCAT 585

## RESULT 104

US-08-544-332-32  
; Sequence 32, Application US/08544332  
; Patent No. 5935777  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Hall, Richard L.  
; APPLICANT: Grudl, Michael E.  
; NUMBER OF SEQUENCES: 77  
; TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Gerard H. Bencen  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/544,332  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/991,867  
; FILING DATE: 07-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/107,755  
; FILING DATE: 19-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/14818  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,685  
; FILING DATE: 30-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/657,584  
; FILING DATE: 19-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bencen, Gerard H.

; REGISTRATION NUMBER: 35,746  
; REFERENCE/DOCKET NUMBER: UF114.C4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 660 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; US-08-544-332-32  
Alignment Scores:  
Pred. No.: 0.0327 Length: 660  
Score: 99.00 Matches: 58  
Percent Similarity: 36.7% Conservative: 32  
Best Local Similarity: 23.7% Mismatches: 75  
Query Match: 6.3% Indels: 80  
DB: 2 Gaps: 11  
US-10-734-719-9 (1-291) x US-08-544-332-32 (1-660)  
Qy 58 GlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMet 77  
Db 43 AAATATTATAATTTAAAAAATAATAATAACAGAGAT----- 81  
Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGlnAsnPheValLysThrPheTyrAsp 97  
Db 82 -----AATGTTATTATAATATAATATAATAAAAAAATAGTTAAT 120  
Qy 98 TyrPheProAspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAla 117  
Db 121 -----TTAGAGAA----- 129  
Qy 118 TyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCys 137  
Db 130 -----TTGCATATAATATAATATATGATAATAATTTTAAATAAT----- 168  
Qy 138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln 157  
Db 169 -----ATTCCAGAAAAATATTTAAAGCTTTATATATATTTCAATTTA----- 207  
Qy 158 AsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAsp 177  
Db 208 AATATTATAATTTAAATTTTATAACAAAATTAATAACATATATAGAT----- 261  
Qy 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLys----- 195  
Db 262 -----ATATCTTATAACAAAATAGCAATATAAGTAATATT 297  
Qy 196 -----AlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211  
Db 298 ATACTACCACATCTATAGATTTTAAAT-----TGTGAA 333  
Qy 212 CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231  
Db 334 TCATGTAATAATAAGTACATATAATTTTATTATAATTTAGTAAATTTAAAAAATTAATA 393  
Qy 232 IleIleGlnGluLys-----AsnAsnTyrThrLysAspIleLeuPro 246  
Db 394 ATATCTAAAAATAAATTTGGTAACTTTTAATTAATGTTTTCTTATTAGTATAGTTGAGTTA 453  
Qy 247 SerSerGluAla-----TyrGlyLysPheSerLysAsnIleAsnPheLys 261  
Db 454 AATATGAATCAATACAAATAAAAGATTATAATTTATAGAAAAATTAATTTAAAA 513  
Qy 262 LysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuPro 281  
Db 514 AAATTAGATATATCTTCAATGTTTAAAAAATAATATA---CATTTGATAAAATTTCCA 570  
Qy 282 SerAspIleLysHis 286

[illegible]

```
QY 87 GluAsnGluAsnPhe---ValLysThrPheTyrAspTyrPheProAspAlaHisLeuGly 105
Db 478 CTAATATATCAATTAAGTTGATGGGAAATTTAAAGAAATTAGATTTTCAGAAATAGCA 537
QY 106 TyrAspPhePheLysGlnLeuLysGluPhe-----AsnAlaTyrPheLys----- 120
Db 538 TATGATAGTAAAGAAATGAAATATTTATATATCAGAACCGCATATATAAGATATTCATG 597
QY 121 -----PheHisGluLeuTyrPheAsnGln 128
Db 598 TTAATAGAAAAAATTTAGATATATTTAGTTGGGATTTTCATCTT---TTTAATGAA 654
QY 129 ArgileThrSer-----GlyValTyrMetCysAlaValAlaIle 141
Db 655 AGTATAGGGAGTAAATACAAATTAATAATATTTAGTACTTATAAATGTTTAAATAGCGCTA 714
QY 142 AlaLeuGlyTyrLysGluLeuSerGlyIleAspPheTyrGlnAsnGlySerSer 161
Db 715 GCTAGTGGC-----GAAATGATACACCAATAAGATAAAATTTAATAACGATTACGCA 768
QY 162 TyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp 181
Db 769 -----ACTAAAAATAAAGAAATCTTAGAA----- 792
QY 182 ArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe----- 199
Db 793 -----TCTTAAACACAGAGATTTAAATTTAATTTGAAATATTTCAT 831
QY 200 -----LeuGluLysThrTyrLysLysLysLeuTyrCysLeuCysProAsn 214
Db 832 GAATTTCAAATTCGTTGGATAAAGAAATTTAAGATA----- 867
QY 215 SerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe----- 231
Db 868 TCCACCCCGTTAATCTTAAAGAAATTTAGAGAGGAGTAAACCTACTTTTATTATAT 927
QY 232 -----IleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuLeuProSer 247
Db 928 AAACAAAACTAATATTGGAAGAAAGCGT-----TTAGAAAGTTTACTAATCAAAACA 981
QY 248 Ser-----GluAlaTyrGlyLys----- 253
Db 982 TCTTCCTCGATTAGAGATATAAAGGAAAAAAATTCACCTTAATTCAGAGAAAAA 1041
QY 254 -----PheSerLysAsnIleAsnPheLysLys-----IleLys 264
Db 1042 GTAACAAAACTATTTTACTCAAAAATCTTTACACACAAATATCAATGTTATTATCAA 1101
QY 265 IleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu-----Pro 281
Db 1102 GTAAAACTCAGACTATGAAAAATTTACAAGATGATGAATTTAAACTAAAAATTTAATCCT 1161
QY 282 SerAspIle 284
Db 1162 GAAGATGTA 1170

RESULT 107
US-09-248-796A-5845
; Sequence 5845, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

```
; SEQ ID NO 5845
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Candida albicans
; NAME/KEY: unsure
; LOCATION: (889), (917), (953), (958), (969), (1034), (1112), (1117)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-248-796A-5845

Alignment Scores:
Pred. No.: 0.0818 Length: 1143
Score: 98.50 Matches: 79
Percent Similarity: 38.0% Conservative: 41
Best Local Similarity: 25.0% Mismatches: 106
Query Match: 6.3% Indels: 90
DB: 3 Gaps: 18

US-10-734-719-9 (1-291) x US-09-248-796A-5845 (1-1143)
QY 16 IleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPhe 35
Db 10 ATGTGTGCTGCCAAATTAATCAAAATTTTGAAGAGTACAAAGTGTACTGAAGATTGTAC 69
QY 36 GluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhe 55
Db 70 -----TATATTAAAGT-----TATATTGCAATTTCGACT 99
QY 56 PheGluGlnTyrTyrThrLeuLysHisLeu-----IleGlnAsnGlnGluTyr 71
Db 100 CAAGACCCAGATGTTGTTTTCAAATACATTCAGCTGCAGGTAGAATGAACCAACCAAA 159
QY 72 GluThrGluLeuIleMetCysSerAsn-----TyrAsnGlnAlaHisLeuGluAsnGlu 89
Db 160 GAAATCGAAGAGTGGTTAGAGATAACAATGTTTATAACGGTGAAAGTG-----AAG 213
QY 90 AsnPheValLysThrPhe-----TyrAspTyr 98
Db 214 AACTTTTGAAGAAGATTTAAATTTGGATGATCAATTAACCTTTAATCATGTTTGTGATAGA 273
QY 99 PheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyr 118
Db 274 TTCAACTTTGTTTCAC-----GATTTGATTTGTTATTGTTACAAAG---AATCAATAC 321
QY 119 PheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSer----- 132
Db 322 TTTAAGTTTATTAAGATTTATGTTCAATCAGTTTAATCCAGCCCAATCTCCACAGATTGTT 381
QY 133 -GlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle---TyrLeuSe 151
Db 382 GCCGGTATTGATGTTGATGTTGATGAAACATCATTAAGGCTGTTAATGCTGTT 441
QY 151 rGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLe 171
Db 442 TTGGGTAGAGTTC-----CAATCAAGAAGATTG 468
QY 171 uLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAs 191
Db 469 GTTGAAGAAGTTCG-----AAAAGAGA 489
QY 191 nThrAspIleLysAla--LeuGluPheLeuGluLysThrTyrLys----- 205
Db 490 AACACAGATTGAAATCTGTTTACCAATTTCTTGAAAAAACATTAGAAGCGGATCTAATGAT 549
QY 206 IleLysLeuTyrCysLeuCysProAsnSerLeuAlaAsnPheIleGluLeuAlaPro 225
Db 550 CAAGAAGTATAC-----AACACTTAGCTAAATCTACATTGAT-----TCC 591
QY 226 AsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIle 245
Db 592 AACAAATTCACAGAAAAAGTTTTCGCAAGAAAAACAACAACACTAC-----GATACATTGTT 645
QY 246 ProSerSerGlu-----AlaTyrGlyLysPheSerLysAsn 257
```







2545 ATGGTCGCTAAATAAACCAGCATGTTTGTGTAATACTAATCCCTTTTAGAGCAATG 2486  
60 TyrThrLeuLeuHisLeuLeuGlnAsnGlnGluThrGluLeuLeuMetCysSer 79  
2485 TAT-----AAGGAGAGGCTTAGTCATAAAGGCTATTCTGATAAAATGCTTTAAGTATT 2432  
80 Asn-----TyrAsnGlnAlaHisLeuGluAsnGluAsnPheVal----- 92  
2431 AATGATAAGAAATATATACTCTAAGGATATTGAAATATCTTGATGTCGCAAGGG 2372  
93 -----LysThrPheTyr-----AspTyrPheProAspAlaHis 103  
2371 GATTATTTAAAAACGTCACGTGGGAGTTTGTTCGCATCTTTTCTGCTCTCAT 2312  
104 LeuGlyTyrAspPhePheGlyGlnLeuLeuGluPheAsnAlaTyrPheLysPheHisGlu 123  
2311 ATT-----AAAGAAGCAATGAGGCTTA----- 2291  
124 IleTyr-----PheAsnGlnAlaGlnThrSerGlyValTyrMetCysAlaValAlaAla 142  
2290 ATTACGAATTTGTCACGAAAGTTGATAATAAAGCCGAGGTGTCGATAAGATTAAATCT 2231  
143 LeuGlyTyrLysGlu-----IleTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159  
2230 CTGGCAAAAAAGCAACAACATGGGATTTAGTACTAAGACACATTACCAAAATGAA 2171  
160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys 179  
2170 AATAATCAAGTTATTGTGAGTAGATCCTTAACTCTACACGGGGCTACTTCGAATGAT 2111  
180 AsnAspArg-----SerHisTyrIleGly----- 187  
2110 AAGCAAAAAAACAAGGTTTCTTCAGAGAGATTAAACGCTTGATAATATCTCGATGATTA 2051  
188 HisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLys 207  
2050 CATTTTGATAAATCTCCCTGGTGAGATTAAACGTTT-----GATAATTAATCTGTAATA 1997  
208 LeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeu 227  
1996 CTTCGCGAGCTGATTAATAAATAAAT-----CCTATTATT 1961  
228 AsnSerAsnPheIleLeuGlnGluLysAsnAsnTyrThrLysAspIleLeuLeuProSer 247  
1960 AATACAAACATCAATTA-----TCATCGTTGTCTAAGGATGTTTA-----AAT 1916  
248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIle----- 263  
1915 TCATTAAATAATGTAGTTT---AAGAAATGTAATTTTTCAGGAGTGATAACAGCCCT 1859  
264 -----LysIleLysGluAsnValTyrTyrLysLeuIleLys----- 275  
1858 AACCTTGAAGGTCGGTTTGTGAAATGCTATTTCGATGATGTTAATTCAACAACATA 1799  
276 AspLeuLeuArgLeuProSerAsp-----IleLysHis 286  
1798 CAGCTTTATGAGACTAATGATGATCTGTGGAATCAGGAATAAAAAACCAATAAAGGT 1739  
287 TyrPheLysGly 290  
1738 ATGTTTAAAGGT 1727

RESULT 111

US-10-114-170-212/c  
; Sequence 212, Application US/10114170  
; Patent No. 685814  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; Burland, Valerie  
; Perna, Nicole T.  
; Plunkett, Guy  
; Welch, Rod

TITLE OF INVENTION: No. 685814el Sequences of E. coli O157

NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Quarles & Brady  
STREET: 1 South Pinckney Street  
CITY: Madison  
STATE: WI  
COUNTRY: US  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/114,170  
FILING DATE: 01-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/453,702  
FILING DATE: 03-DEC-1999  
APPLICATION NUMBER: 60/110,955  
FILING DATE: 04-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95017  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 251-5000  
TELEFAX: (608) 251-9166  
INFORMATION FOR SEQ ID NO: 212:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7425  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 212:  
US-10-114-170-212  
Alignment Scores:  
Pred. No.: 1.35 Length: 7425  
Score: 98.00 Matches: 82  
Percent Similarity: 37.4% Conservative: 54  
Best Local Similarity: 22.5% Mismatches: 122  
Query Match: 6.3% Indels: 107  
DB: 3 Gaps: 21  
US-10-734-719-9 (1-291) x US-10-114-170-212 (1-7425)  
Qy 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLys-----GluIleAsp 17  
Db 2724 AAAAAATTCGCAATCCAGGTCGACACCCCTGGGTTAAAAATGGAATAATAGCTATTTT 2665  
Qy 18 TyrSer-----ArgLeu 21  
Db 2664 TACAGTATTCACCGTAAATATAAGTAACACTTGTGCTAACCTAATACAGTACAGTTA 2605  
Qy 22 ProAsnAspPheAsp-----ValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 39  
Db 2604 ATTAACACATAGAGATTATTGTTTTTCAACAACACTCAATATTACACGCCGGAGGA-TTT 2546  
Qy 40 TyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 59  
Db 2545 ATGGTCGCTAAATAAACCAGATGTTTTTGTAACTAATAATCCTTTTGTAGAACCAATG 2486  
Qy 60 TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSer 79  
Db 2485 TAT-----AAGGAGAGGCTTAGTCATAAAGGCTATTCTGATAAAATGCTTTAAGTATT 2432  
Qy 80 Asn-----TyrAsnGlnAlaHisLeuGluAsnGluAsnPheVal----- 92  
Db 2431 AATGATAAGAAATATATACTCTAAGGATATTGAAATATCTTGATGTCGCAAGGG 2372



```
Db 714 TCATCCAGTAGTAGCTGCTTACAGTACACACATTAAATAATTGATGATCATAGAT 655
Qy 155 -----PheTyrGlnAsnGlySerTyr----- 162
Db 654 TCATCGCCATTTAAATGATATATAAAGCGATGTTTGAAATTTATCTCGATATCGTGCA 595
Qy 163 ---AlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp 181
Db 594 ATTATCGTTTCTACTAGCAGCAGCAAAATTA-----GATGTTGAAACACGT 550
Qy 182 ArgSerHisTyrIle-----GlyHisSerLysAsnThrAspIle 194
Db 549 ATTAACCATACTATTCAGTTGTTAATATTCCTGCTGCTATAGTGAAACAATTGATACA 490
Qy 195 LysAlaLeuGluPheLeuGluLysThrTyrLys---IleLysLeuTyrCysLeuCysPro 213
Db 489 CCAGTCCAAACATTAGATCAGCGCAGTGTTAAATTAATTCCTGCTGCTGTTTACCT 430
Qy 214 AsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIle 233
Db 429 GAAAAACAATTACATCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 382
Qy 234 GlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGlu----- 249
Db 381 -----TATGTCCTCCAAAAATAGAGCTTCATATGTAT 352
Qy 250 AlaTyrGlyLysPheSerLysAsnIleAsn-----PheLysLysIleLysIleLysGlu 267
Db 351 GCTTTGGTCTCTGAAAGTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 292
Qy 268 AsnValTyrTyrLys 272
Db 291 CATGTTTACTTAGCG 277

RESULT 113
US-08-781-986A-109/c
; Sequence 109, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17846 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: linear
US-08-781-986A-109
Alignment Scores:
Pred. No.: 5.38 Length: 17846
Score: 97.50 Matches: 65
Percent Similarity: 35.7% Conservative: 51
Best Local Similarity: 20.0% Mismatches: 114
Query Match: 6.2% Indels: 95
Db: 3 Gaps: 16

US-10-734-719-9 (1-291) x US-08-781-986A-109 (1-17846)
Qy 12 SerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsn 31
Db 1158 ACTTTAAACATGTTGAAATAGT-----CATGACTTTAGAAATATATGATCAAGAG 1108
Qy 32 GlnPhe-----TyrPheGluAspLysTyrTyr----- 40
Db 1107 CGTTTCTTAATGATGATGCCCATTTTCAAGATCCTAAATATCGTATATTAGATTATGTAAT 1048
Qy 41 -----LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePhe 56
Db 1047 CATTTTGATAGTCAAGACGAAAGTAAACGCTGATTTTATGATGTCGCTGCTTTT 988
Qy 57 -----GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGln 69
Db 987 AGTTGTAGTAGATTCCTGTTGATAAAACAACAACATTTATGTAATTTTCTATATATCCA 928
Qy 70 GluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGlu 89
Db 927 GAAGTGACACAAATAGAAAAGTATTTTCTTACAGAGTGGCAACACAGAGTACAA 868
Qy 90 AsnPheValLysThrPhe-----TyrAspTyrPheProAspAlaHisLeuGly 105
Db 867 AAAATATTGTTTACTATGCTAAACAACAATATTTCTTTAATAATAACGGAATTAGGC 808
Qy 106 TyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyr 125
Db 807 GCATTTTTCATTAACAACATA-----TATCAACATCGGGATTTATTT 766
Qy 126 PheAsnGlnArgIleThrSerGlyValTyrMetCys----- 137
Db 765 TTCAGTGATCGT-----AATGTGTACACAGCACCTATTTTCAACTTAACGCTGAG 715
Qy 138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp----- 154
Db 714 TCAATCCAGTAGTAGCTGCTGTCCACAGTACACACATTAATAATATTGATGCATTAGAT 655
Qy 155 -----PheTyrGlnAsnGlySerTyr----- 162
Db 654 TCATCGCCATTTAAATAATGATATATAAAGCGATGTTTGAAATTTATCTCGATATCGTGCA 595
Qy 163 ---AlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAsp 181
Db 594 ATTATCGTTTCTACTAGCAGCAGCAAAATTA-----GATGTTGAAACACGT 550
Qy 182 ArgSerHisTyrIle-----GlyHisSerLysAsnThrAspIle 194
Db 549 ATTAACCATACTATTCAGTTGTTAATATTCCTGCTGCTATAGTGAAACAATTGATACA 490
Qy 195 LysAlaLeuGluPheLeuGluLysThrTyrLys---IleLysLeuTyrCysLeuCysPro 213
Db 489 CCAGTCCAAACATTAGATCAGCGCAGTGTTAATAATTAATTTCTGCTGCTGTTTACCT 430
Qy 214 AsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIle 233
Db 429 GAAAAACAATTACATCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 382
Qy 234 GlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGlu----- 349
Db 381 -----TATGTCCTCCAAAAATAGAGCTTCATATGTAT 352
```

```
Qy 250 AlaTyrGlyLysPheSerLysAsnIleAsn-----PheLysLysIleLysIleLysGlu 267
Db 351 GGTITGGTCTGAAAGTAAATAATTAATGAATTGATACAAAATAATGGAATGGAATAAT 292

Qy 268 AsnValTyrTyrLys 272
Db 291 CATGTTACTTACGC 277

RESULT 114
US-09-710-279-655
; Sequence 655, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 655
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-655

Alignment Scores:
Pred. No.: 0.129 Length: 1188
Score: 97.00 Matches: 62
Percent Similarity: 36.3% Conservative: 33
Best Local Similarity: 23.7% Mismatches: 87
Query Match: 6.2% Indels: 80
DB: 3 Gaps: 13

US-10-734-719-9 (1-291) x US-09-710-279-655 (1-1188)

Qy 57 GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIle 76
Db 436 GAGAACTATCAAAATGCTAAAGCAAAATATATAAAATCAAACTAAAGATGATTATCTC 495

Qy 77 MetCysSerAsnTyrAsnGlnAlaHisLeu---GluAsnGluAsnPhe---ValLysThr 94
Db 496 ATTTGT---AATTATCATCAAGACACCTAATTGAATCAGAAAATCTAGAACGGAACA 552

Qy 95 PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu 114
Db 553 TTTTAT-----TTTTCACACACGCAAGAAGTT 579

Qy 115 PheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn----- 127
Db 580 GATGGATATACATTAAAGATGGTTTCATTGTTTAAACGGCATTCGCATTATTAACTACT 639
Qy 128 -----Gln 128

Db 640 AAAGACTTAGTCTACCAGGAACATAACTCTGGAATAATTTTAGCAGCTCTTAGCA 699

Qy 129 ArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148
Db 700 TCAATCATTTGGAGTCCAGTCAAAAGCTATTGTAGATAGCTTT-----GTT 747

Qy 149 TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGln 168
Db 748 ACTTTTCCGGTATTGATCAT----- 768

Qy 169 GluAsnLeuLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis 188
Db 769 -----AGACTTCAGTATATTGGTACAAATCCACAAATAATATTATATGAT 816
```

```
Qy 189 SerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeu 208
Db 817 TCAGAAAGCAACTAAATACTTTAGCTACTCAATTTGCGCTTAACCTCTTTTGATCAACCAATT 876

Qy 209 TyrCysLeuCysProAsnSerLeuLeuAlaAsn---PheIleGluLeuAlaProAsnLeu 227
Db 877 ATTTGGTTGTGTGGTGGATTAGATCGTGGTAATGAATTCGATGAACATTATCCTTATATG 936

Qy 228 AsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSer 247
Db 937 GAAAT-----GTACGTGTGATGGTTGTTT 963

Qy 248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLys---LysIleLysIleLys 266
Db 964 CGAGAAACACAAAGATAAATTTGCTAAATTGGGAATAGTCAAGGTAAGTATGTGATTAAA 1023

Qy 267 -----GluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSer 282
Db 1024 GCAACAGATGTAGAGGATGCTGTTGATAAA---ATTCAAGATATAGTCGAG---CCAAAT 1077

Qy 283 AspIle 284
Db 1078 GATGTT 1083

RESULT 115
US-09-134-001C-1001
; Sequence 1001, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1001
; LENGTH: 1374
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1001

Alignment Scores:
Pred. No.: 0.159 Length: 1374
Score: 97.00 Matches: 62
Percent Similarity: 36.3% Conservative: 33
Best Local Similarity: 23.7% Mismatches: 87
Query Match: 6.2% Indels: 80
DB: 3 Gaps: 13

US-10-734-719-9 (1-291) x US-09-134-001C-1001 (1-1374)

Qy 57 GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIle 76
Db 622 GAGAACTATCAAAATGCTAAAGCAAAATATATAAAATCAAACTAAAGACGATTATCTC 681

Qy 77 MetCysSerAsnTyrAsnGlnAlaHisLeu---GluAsnGluAsnPhe---ValLysThr 94
Db 682 ATTTGT---AATTATCATCAAGGCACCTAATTGAATCAGAAAATCTAGAACGGAACA 738

Qy 95 PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu 114
Db 739 TTTTAT-----TTTTCACACACGCAAGAAGTT 765

Qy 115 PheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn----- 127
Db 766 GATGGATATACATTAAAGATGGTTTTCATTGTTTAAACGGCATTCGCATTATTAACTACT 825

Qy 128 -----Gln 128
```

```
Db      826 AAAGACTTAGTGTACCGGAGAACATAACCTGGAAAAATATTTTACAGCTGTTCTAGCA 885
Qy      129 ArgileThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148
Db      886 TCATCATCTGCTGGAGTGCAGTCAAAGCTATTGTAGATAGTCTT-----GTT 933
Qy      149 TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGln 168
Db      934 ACTTTTCCGGTATTGATCAT----- 954
Qy      169 GluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis 188
Db      955 -----AGACTTCAGTATATTGGTACAAATCGCACAAATAAATAATATTATATGAT 1002
Qy      189 SerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeu 208
Db      1003 TCAGAGCAACTAATCTACTTCTAGCTACTCAATTTCCGCTTAACTTTTGTATCAACCAATT 1062
Qy      209 TyrCysLeuCysProAsnSerLeuLeuAlaAsn---PheIleGluLeuAlaProAsnLeu 227
Db      1063 ATTTGGTTGTGCTGGATAGATCGTGGTAAATTCGATGAACCTATTTCCTTATATG 1122
Qy      228 AsnSerAsnPhelIleGlnGluLysAsnAsnTyrThrLysAspIleLeuLeuProSer 247
Db      1123 GAAAAAT-----GTACGTGTGATGGTTGTTTTT 1149
Qy      248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLys---LysIleLysIleLys 266
Db      1150 GGAGAACACACAGATAAATTTGCTAAATTTGGAAATAGTCAAGTAAGTATGTGATTAA 1209
Qy      267 -----GluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSer 282
Db      1210 GCAACAGATGTAGAGGATGCTGTTGATAAA---ATTCAAGATATAGTCGAG---CCCAAT 1263
Qy      283 AspIle 284
Db      1264 GATGTT 1269

RESULT 116
US-09-710-279-3574/C
; Sequence 3574, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3574
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3574

Alignment Scores:
Pred. No.: 0.563 Length: 3334
Score: 97.00 Matches: 62
Percent Similarity: 36.3% Conservative: 33
Best Local Similarity: 23.7% Mismatches: 87
Query Match: 6.2% Indels: 80
Db: 3 Gaps: 13

US-10-734-719-9 (1-291) x US-09-710-279-3574 (1-3334)
Qy      57 GluGlnTyrThrLysLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIle 76
```

```
Db      695 GAGAACTATCAAAATGCTAAAAACCAATATATAAAATCAAACTAAAGATGATTATCTC 636
Qy      77 MetCysSerAsnTyrAsnGlnAlaHisIleu---GluAsnGluAsnPhe---VallyThr 94
Db      635 ATTGT---AATTATCATCAAGACACACTAATTGAATCAGAAAAATCTAGAGCGAAAAACA 579
Qy      95 PheTyrAspTyrPheProAspAlaHisIleuGlyTyrAspPheLysGlnLeuLysGlu 114
Db      578 TTTTAT-----TTTTCACACACAGCAAGATT 552
Qy      115 PheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn----- 127
Db      551 GATCGGATATACATTAAAGATGGTTTCATTGTTTAAACGCATTTCGCATTATTAACT 492
Qy      128 -----Gln 128
Db      491 AAAGACTTAGTGTACCGAGAGAACATAACCTGGAAAAATATTTTACAGCTGTTCTAGCA 432
Qy      129 ArgileThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148
Db      431 TCAATCATCTGCTGGAGTCCAGTCAAAGCTAATTGTAGATAGTCTT-----GTT 384
Qy      149 TyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGln 168
Db      383 ACTTTTCCGGTATTGATCAT----- 363
Qy      169 GluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis 188
Db      362 -----AGACTTCAGTATATTGGTACAAATCGCACAAATAAATAATATTATATGAT 315
Qy      189 SerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeu 208
Db      314 TCAAAAGCAACTAATCTACTTCTAGCTACTCAATTTGCGCTTAACCTTTTGATCAACCAAT 255
Qy      209 TyrCysLeuCysProAsnSerLeuLeuAlaAsn---PheIleGluLeuAlaProAsnLeu 227
Db      254 ATTTGGTTGTGCTGGATAGATCGTGGTAAATTCGATGAACCTATTTCCTTATATG 195
Qy      228 AsnSerAsnPhelIleGlnGluLysAsnAsnTyrThrLysAspIleLeuLeuProSer 247
Db      194 GAAAAAT-----GTACGTGTGATGGTTGTTTTT 168
Qy      248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLys---LysIleLysIleLys 266
Db      167 GGAGAACACACAGATAAATTTGCTAAATTTGGAAATAGTCAAGTAAGTATGTGATTAA 108
Qy      267 -----GluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSer 282
Db      107 GCAACAGATGTAGAGGATGCTGTTGATAAA---ATTCAAGATATAGTCGAG---CCCAAT 54
Qy      283 AspIle 284
Db      53 GATGTT 48

RESULT 117
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-3-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Alignment Scores:  
Pred. No.: 4.52e+03 Length: 1830121  
Score: 97.00 Matches: 61  
Percent Similarity: 37.7% Conservative: 58  
Best Local Similarity: 19.3% Mismatches: 127  
Query Match: 6.2% Indels: 70  
DB: 3 Gaps: 15

US-10-734-719-9 (1-291) x US-09-557-884-1 (1-1830121)

QY 39 TyrTyrLeuGlyLysLysCys-----LysAlaValPheTyrThrProAsnPhePhePhe 56  
Db 1292100 TGGTATATGGCGCGCGTAGTCGGAAGAAATCAAGATGACATAAGTAATAAATCTGCC 1292159  
QY 57 GluGlnTyrTyrThr---LeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeu 75  
Db 1292160 CGAGATTATGTCGGGAGTGAATTTTACTTTCCATCAACTGATGAAGCTGTAGAT 1292219  
QY 76 IleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPhe 95  
Db 1292220 TTGTTTTTAGATATGTTCCAAAGCAAGAAATCGAAATGAA-----ATCGAAAGTCAC 1292273  
QY 96 TyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheGlnLeuLysGluPhe 115  
Db 1292274 TCCCAATTTGAAGCAGAACTTACACTTGGC---AATCTGTTTCGTTCTCGAGGCGAAGTT 1292330  
QY 116 AsnAlaTyrPheLysPheHisGluLeuTyr-----PheAsnGln 128  
Db 1292331 GATCGTCATTCGCAATTCATCACTTAGATCTAGCTCTAATTAATCTTTTGAACAA 1292390  
QY 129 ArgIleThrSerGlyValTyrMetCysAlaValAlaLeuLeuGlyTyr----- 145  
Db 1292391 AAATTAATTCGCAAAACACAGCTTCGAGATTTTATGGTGGTGGTTTTTTGACCGA 1292450  
QY 146 ---LysGluLeuTyrLeuSerGlyIleAsp-----PheTyrGlnAsnGlySerSer 161  
Db 1292451 GCTGAGAATCTTTATATTACTAGTAGTAGTAACCTGAGTTTGCAGAAAAATGCTTTACAG 1292510  
QY 162 TyrAlaPhe-----AspThrLysGln-----Glu 169  
Db 1292511 CAATTTTAGTATTATCAAAAGCAAGAAAGTGAAGAAAAAGCAGTCAATATTGCTGAA 1292570  
QY 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArg-----SerHisTyrIleGly 187  
Db 1292571 AAATCGCTAAATAAAGCCGCAAGAGAAATATATTGAACTTGCCTCAATGTTATTGTGAA 1292630

QY 188 HisSerLysAsnThrAsp-----IleLysAlaLeuGluPheLeuGluLysThr 203  
Db 1292631 TATTCAAAAGTTTAGAGCCTGAAGTCGGTTGAAAAACGCACTGTTTACAAAAGCT 1292690  
QY 204 TyrLysIleLeuLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223  
Db 1292691 TTGTCGGTTTCTCCGACTTGTGTTCGGCTTCTTTATTGCTTGCAAAATTTAGCAATGCTT 1292750  
QY 224 -----AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsn 237  
Db 1292751 GACGGTCAATATCAACAGCAGTAAATAATTAGAAAT---GTTTTGGAGCAAAATCCA 1292807  
QY 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257  
Db 1292808 GATTATACGGGTGAAATTTTGTACCTTTAAACATTGTTACGAAGAATTAAATCAGTTA 1292867  
QY 258 IleAsnPheLysLysIleLysIle----- 265  
Db 1292868 GATAATTTTGAGTTATTTTAAATTCGTGAGGTCAAAATTATTAATGATGAAGTAGAG 1292927  
QY 266 -----LysGluAsnValTyrTyrLysLeu 273  
Db 1292928 CTGGCAGTAGCTAAATAATTGAAGAAAAAGATGGCAAGTCTGCTGCTCAAGCTAAGCTT 1292987  
QY 274 IleLysAspLeuLeuArgLeuProSerAsp---IleLysHisTyrPhe 288  
Db 1292988 TATCAGCAACTCAGGAAAGCCCAAGTACTTTAATTTTCCATCGTTTT 1293035  
RESULT 118  
US-09-643-990A-1  
Sequence 1, Application US/09643990A  
Patent No. 6528289  
GENERAL INFORMATION:  
APPLICANT: Robert D. Fleischmann  
Mark D. Adams  
Owen White  
Hamilton O. Smith  
J. Craig Venter  
TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB186P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs

		TYPE: nucleic acid		STRANDEDNESS: double		TOPOLOGY: linear		SEQUENCE DESCRIPTION: SEQ ID NO: 1:		US-09-643-990A-1	
		Alignment Scores:		Pred. No.:		Score:		Percent Similarity:		Best Local Similarity:	
		Query Match:		Indels:		Gaps:		Length:		Matches:	
		DB:		61		58		127		70	
				15							
US-10-734-719-9 (1-291) x US-09-643-990A-1 (1-1830121)											
QY	39	TyTyrLeuGlyLysLysCys	-----LysAlaValPheTyrThrProAsnPhePhe	56							
Db	1292100	TGGTATATGGGGCGGTAGTCGAAAGAAAGATCAAGATGACATAGTAAGTAATAAACTCTCC	1292159								
QY	57	GluGlnTyrTyrThr	-----LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeu	75							
Db	1292160	CGAGATTATGTCGGGAGTGAATTTTACTTCCCAATCAACTGATAAAGCTGTAGAT	1292219								
QY	76	IleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPhe	95								
Db	1292220	TGCTTTTATAGATATGTTGCAAAAGCAAGAAATCGAAATGAA	1292273								
QY	96	TyrAspTyrPheProAspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPhe	115								
Db	1292274	TCCCAATTGAAGCAGAACTTACACTTGGC	1292330								
QY	116	AsnAlaTyrPheLysPheHisGluIleTyr	-----PheAsnGln	128							
Db	1292331	GATCGTGCATTGGCGATTCACTAGCTTTAGATCTTCTAATATTACTTTTGAACAA	1292390								
QY	129	ArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyr	145								
Db	1292391	AAATTACTTGCAAAACACAGCTTGCTCGAGATTTATGGTGTGGTTTTTTTGAACCGA	1292450								
QY	146	-----LysGluIleTyrLeuSerGlyIleAsp	-----PheTyrGlnAsnGlySerSer	161							
Db	1292451	GCTGAGAACTCTTTATATTACTAGTAGAGAACCTGAGTTTGCAGAAATGTCTTACAG	1292510								
QY	162	TyrAlaPhe	-----AspThrLysGln	169							
Db	1292511	CAACTTTTAGTGATTATCAAAAGACAAAGAAATGGAAGAAAGCAGTCAATATTGCTGAA	1292570								
QY	170	AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArg	-----SerHisTyrIleGly	187							
Db	1292571	AAACTCGCTAAATAAAGCCGCAAGAGAATAATATTGAACCTGGCGCAATGTTATTGTGAA	1292630								
QY	188	HisSerLysAsnThrAsp	-----IleLysAlaLeuGluPheLeuGluLysThr	203							
Db	1292631	TATTACAAAGTTTAGAGCCCTGAAAGTCGGCTGAAAAACGACGAGTGTTTTACAAAAGGCT	1292690								
QY	204	TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu	223								
Db	1292691	TTGTGCGTTTCTCCGACTGTGTGCGCCTTCTTTATTGCTTGCAAAATTAGCAATGCTT	1292750								
QY	224	-----AlaProAsnLeuAsnSerAsnPheIleIleGlnGlnLysAsn	237								
Db	1292751	GACGGTCAATATCAACAGCAGTAAATAATTAGAAAT	1292807								
QY	238	AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn	257								
Db	1292808	GATTATACGGGTGAAATTTTGTACCTTTAAACATTTGTTACGAAAGAAATTAATCAGTTA	1292867								
QY	258	IleAsnPheLysLysIleLysIle	-----265								
Db	1292868	GATAATTTTGAGTTATTTTAAATTCGTCAGGCTCAATATTAAATGATGAAGTAGAG	1292927								
QY	266	-----LysGluAsnValTyrTyrLysLeu	273								



```
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
```

```
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
```

```
Alignment Scores:
Pred. No.: 4.52e+03 Length: 1830121
Score: 97.00 Matches: 61
Percent Similarity: 37.7% Conservatives: 58
Best Local Similarity: 19.3% Mismatches: 127
Query Match: 6.2% Indels: 70
DB: 3 Gaps: 15
```

US-10-734-719-9 (1-291) x US-10-158-865-1 (1-1830121)

QY 39 TyrTyrLeuGlyLysLysCys-----LysAlaValPheTyrThrProAsnPhePhePhe 56  
Db 1292100 TGTATATGGGGCGGTAGTCGAGAGAGATCAAGATGACATAAGTATAAACTCTCC 1292159  
QY 57 GluGlnTyrThr---LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeu 75  
Db 1292160 CGAGATTATGTTACGGAGGTAAATTTTATCTTCCAACTCAAACATGATAAAGCTGTAGAT 1292219  
QY 76 IleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPhe 95  
Db 1292220 TTGTTTTAGATATGTTGCAAGCAAGCAAGAAATCGAAATGAA-----ATCGAAAGTCAC 1292273  
QY 96 TyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPhe 115  
Db 1292274 TCCCAATTTGAAGCAGAACTTACACTTGGC---AATCTGTTTCTGCTTCGAGGCGAAGTT 1292330  
QY 116 AsnAlaTyrPheLysPheHisGluIleTyr-----PheAsnGln 128  
Db 1292331 GATCGTGCATTTGCGCATTCATCAAGCTTTAGATCTTCTAGTCTTAATTAATCTTTTGAACAA 1292390  
QY 129 ArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyr----- 145  
Db 1292391 AAATTAATCTGCAAAACACACAGCTTCTCGAGATTTTATGGTGTGTTTTTTTGACCGA 1292450  
QY 146 ---LysGluIleTyrLeuSerGlyIleAsp-----PheTyrGlnAsnGlySerSer 161  
Db 1292451 GCTGAGAATCTTTATATTTTACTAGTAGTGAACCTGAGTTTGCAGAAAATGCTTTACAG 1292510  
QY 162 TyrAlaPhe-----AspThrLysGln-----Glu 169  
Db 1292511 CAACCTTTTAGTGATTTATCAAAAGACAAAGAAATGSGAAAAAGCAGTCAATATTGCTGAA 1292570  
QY 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArg-----SerHisTyrIleGly 187  
Db 1292571 AAATCGCTAAATTAAGCCCAAGAGATTAATATTGAACCTTGGCCATGTTATGTGAA 1292630  
QY 188 HisSerLysAsnThrAsp-----IleLysAlaLeuGluPheLeuGluLysThr 203  
Db 1292631 TATTCACAAAGTTTAGAGCTGAAAGTCGGTTCGAAAGCAGTGTGTTTACAAAAGGCT 1292690  
QY 204 TyrLysIleLysLeuTyrCysLysCysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223  
Db 1292691 TTGTCGGTTCCTCCGACTTGTGTTGCGCTTCTTTTATGCTTGCAAAATTTAGCAATGCTT 1292750  
QY 224 -----AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsn 237  
Db 1292751 GACGGTCAATATCAACAGCAGTAAATAATATTAGAAAT-----GTTTTGGCAAAATCCA 1292807  
QY 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257  
Db 1292808 GATTATACGGGTGAAATTTGTTACTTTAAACATGTTTACGAAGAATTAATCAAGTTA 1292867  
QY 258 IleAsnPheLysLysIleLysIle----- 265  
Db 1292868 GATAATTTTGATTTATTTTAAATTCGTGCGAGTCAAATTTATTAATGATCAAGTAGAG 1292927  
QY 266 -----LysGluAsnValTyrTyrLysLeu 273  
Db 1292928 CTGCACCTAGCTAAATTAATTGAAGAAAAGATGCGAGTCTGCTGCTCAAGCTT 1292987  
QY 274 IleLysAspLeuLeuArgLeuProSerAsp---IleLysHisTyrPhe 288  
Db 1292988 TATCAGCAACTCACGAAAAGCAAGTACTTAAATTTTCCATCGTTTT 1293035

RESULT 120  
US-09-662-254B-26  
; Sequence 26, Application US/09662254B  
; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Allison Louise

; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UP-221C1XC1  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-662-254B-26  
  
Alignment Scores:  
Pred. No.: 30.6 Length: 50000  
Score: 96.50 Matches: 52  
Percent Similarity: 32.0% Conservative: 29  
Best Local Similarity: 20.6% Mismatches: 73  
Query Match: 6.2% Indels: 99  
DB: 3 Gaps: 10  
  
US-10-734-719-9 (1-291) x US-09-662-254B-26 (1-50000)  
QY 56 PheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeu 75  
Db 47793 TATCATGATATAAATTTTATCAATATATACATTTTAATAAATAAAACATATAATGTTG 47852  
QY 76 IleMetCys---SerAsnTyrAsnGln----- 83  
Db 47853 ATTAATGTTATGAGTAATAATTAATTTACCTGTTACTAATAATATAGACTTTATAGAGAT 47912  
QY 84 -----AlaHisLeuGluAsnGluAsnPheValLysThrPheTyr 96  
Db 47913 ATTGCAATAATATCAATATATACACATATTCATATAAATAATTTAATTAAGATAATTTTT 47972  
QY 97 AspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsn 116  
Db 47973 AAAAAATTTGACGAT----- 47987  
QY 117 AlaTyrPheLysPheHisGluIleTyrPhe----- 126  
Db 47988 -----TTAAAGAAATAATTTTGTACCGGTAATATAGATATTTTATTGAT 48035  
QY 127 AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys 146  
Db 48036 AATGATATAGTAATAAATAATGAATATATA-----CACAAATTTATCATTATAGA 48083  
QY 147 GluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGly----- 159  
Db 48084 AAAATATTAAAGAAATGTTTAGAAAACAATAGATGATAACGAATATAGATATCATATTTTA 48143  
QY 160 -----SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAsp 177  
Db 48144 AGAGATGAATGTTATGAAATTTGATCATTTTCGACGATATAATAAATAATATATATGCGCAAGT 48203  
QY 178 PheLysAsnAspArgSerHis---TyrIleGlyHisSerLysAsnThrAspIleLysAla 196  
Db 48204 TATTCGGAAGATAAAAAATATAAATATCTAATATTAATAAATAAAGCAAGATATCAGAT 48263  
QY 197 LeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu 216  
Db 48264 TTAAAA----- 48269  
QY 217 LeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys 236  
Db 48270 -----TCATCAAAAGATATAATAAATATAGAT 48296  
QY 237 AsnAsnTyrThrLys---AspIleLeuIleProSerSerGluAlaTyrGlyLys----- 253

Db 48297 AATAATATACCAAAATGTGATATTTTAAATACATCTAGTCTCCTTTTGGTGATGATAAT 48356  
QY 254 -----PheSerLysAsnIleAsnPheLys 261  
Db 48357 GCGTGTGGATATTTATATCAAAAGTTATAAATATTAAA 48395  
RESULT 121  
US-09-861-451A-45  
; Sequence 45, Application US/09861451A  
; Patent No. 6759516  
; GENERAL INFORMATION:  
; APPLICANT: Commonwealth Scientific & Industrial Research Orga  
; TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences  
; FILE REFERENCE: FF34033/01  
; CURRENT APPLICATION NUMBER: US/09/861,451A  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: PP7273  
; PRIOR FILING DATE: 1998-11-20  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 1101  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Clone pAD922  
US-09-861-451A-45  
Alignment Scores:  
Pred. No.: 0.152 Length: 1101  
Score: 96.00 Matches: 74  
Percent Similarity: 34.6% Conservatives: 35  
Best Local Similarity: 23.5% Mismatches: 112  
Query Match: 6.1% Indels: 94  
DB: 3 Gaps: 16  
US-10-734-719-9 (1-291) x US-09-861-451A-45 (1-1101)  
QY 14 LysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPhe 33  
Db 31 AAAACTTTAGACCAAAATATCTGAAATTTGACCTTGAAATGCTGAAAAGGTAACCTT 90  
QY 34 TyrPheGluAsp----- 37  
Db 91 GAAATTCAGATTTTCATTAAATTCCTGCAATCTCGAATCGAATGAAAAAATAGAGAAAATT 150  
QY 38 ---LysTyrTyrLeuGlyLysCysLysAlaValPheTyrThrProAsnPhePhe 56  
Db 151 GCGAAATTTCTACGATTAATATCAGAGGCAAAATC-----AACGAATTTT 201  
QY 57 GluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIle 76  
Db 202 AATCTCTATTGCTTTATTTTGAAAAATTTACAAAA-----TTA 240  
QY 77 MetCysSerAspTyrAsnGlnAlaHisLeuGluAsnGluAsnPheVallysthrPheTyr 96  
Db 241 TTTAGCTCATCATATAAT-----CTTGGCTATGAAATGTGGCCAAA---TTATAT 288  
QY 97 AspTyrPheProAspAlaHisLeuGlyTyr----- 106  
Db 289 GATTATTTCTATGAAGTCCAAAAAATTTACCGACAAAAACAGACGAAAAAGTCGAATTT 348  
QY 107 -----AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119  
Db 349 GACTACCGCAGTGCTAAAAAGATATGAGACACGCTAAAAA-----ATA 396  
QY 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal 139  
Db 397 AAGCAAGAAAAAGCTTTTTCATTAATAAACATTAAT----- 432  
QY 140 AlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159  
Db 433 GTAAAAAGCGCTTAATTTAAAAAAGAGCGCCAACTCGAGATTGACAAATTC----- 483

QY 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLys 179  
Db 484 -----ACCGCTCAAAACAATTTGTTGACTTCCTATATTGACGAATTT 525  
QY 180 AsnAspArgSerHisTyrIleGlyHisSerLys-----AsnThrAspIleLysAla 196  
Db 526 AAT---TATGAATATATAAATTCAAATAACAAAGCGCTAGTAAACAACAGATCTAAAAAAAT 582  
QY 197 LeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu 216  
Db 583 TATTCATTTTAAAAAACAAGCAATAATC-----AATAAGAA 621  
QY 217 LeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys 236  
Db 622 ATTGCCAAATTTCTGAT-----AGGAGAAATATTTTGTACTTGAAAAA 666  
QY 237 AsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLys 256  
Db 667 AACCTTTTTC-----TTTCTTAATATTTCTGAGATTGAAAAATTTATTTGAAAT 717  
QY 257 AsnIleAsnPheLysLysIleLysGluAsnValTyrTyrLysLeuIleLysAsp 276  
Db 718 ATGAATAATTTCAAAAAAGTCAAAATGAAAG-----TATAAAGTTTGTACTTTCGAT 771  
QY 277 LeuLeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 772 -----AAAAAGATGAAAAAATTTATACAAATACAAA 804

## RESULT 122

US-09-248-796A-8070

; Sequence 8070, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 8070

; LENGTH: 1314

; TYPE: DNA

; ORGANISM: Candida albicans

US-09-248-796A-8070

## Alignment Scores:

Pred. No.: 0.196 Length: 1314

Score: 96.00 Matches: 63

Percent Similarity: 38.8% Conservatives: 45

Best Local Similarity: 22.7% Mismatches: 100

Query Match: 6.1% Indels: 70

DB: 3 Gaps: 14

US-10-734-719-9 (1-291) x US-09-248-796A-8070 (1-1314)

QY 60 TyrThrLeuLysHisLeuIleGlnAsnGln-----GluTyrGluThrGluLeuIleMet 77  
Db 454 TATATTCATATGTCCTCTATTGAAATATCATCATATTGAGGGGTTAATTTAATGATT 513  
QY 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsn----- 88  
Db 514 GCTTCAAAATGAAATTTTATACATCTGATCATATTCGAAACAAAATTAGAAGGTAT 573  
QY 89 -----GluAsnPheVallysthrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106  
Db 574 TATATCAAAATATTTTCATTCGAATCTTATATGGA-----GAAGAGGCATTTATATCCAT 627

```
Qy 107 AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPhe 126
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 628 GATGATTATTAGTACTTTGAAAGCTGTGTACATTGGGGGAATATAAAGGGGATTTTATAT 687
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 127 -----AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIle 141
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 688 AAATTGGATGACTATATATAATAATTGATGAAATGAAATAATGTTTGATAATTC 747
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 142 AlaLeuGlyTyrLysGluLeuTyrLeuSerGlyLeuAspPheTyrGlnAsnGlySer 161
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 748 CCTGATCCATGGAAGAAATGAAATTAATGATATATGATCATGATATCTAGTGGTACTTCG 807
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 162 ---TyrAlaPheAspThrLysGlnGluAsnLeu-----LysLeu 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 808 AAATATTCAATATCCGTAAGAAAGTATTGATATTGTGCCGATGCGGAATTTTCAAAATG 867
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 175 AlaProAspPheLysAsn-----AspArgSerHisTyrIleGlyHisSerLysAsn 191
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 868 ATTCCTACAACAATAATTAATGATGATGTTGTTATATATTAATTTATGTCATGAATTGAA 927
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 192 ThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 928 ACTAATCCTATAAATATCATTTACGTTCTAAACT-----AAAGATTTATGTGTT 978
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 212 CysProAsnSerLeu-----LeuAla 218
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 GATCCCAATTAATTATCAGAATTAATAATGAAATTTGGTCAGCAATAAATCAATTAATTA 1038
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 219 AsnPheIleGluLeuAlaProAsnLeuAsnSerAsn-----PheIleIleGln 234
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1039 AATTATATATTTCTTAATTAATTAATTAATTTTAATTAATACCGGTGGTGAATTTTATTACAA 1098
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 235 GluLysAsn-----AsnTyrThrLysAspIleLeuProSerSerGluAlaTyrGly 252
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1099 AATTCGAAATTCCTAAATTAACCTCCTGATTAAAT----- 1134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 253 LysPheSerLysAsnIleAsnPhe-----LysLysIleLysIleLysGluAsnValTyr 270
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1135 -----AATCAAAATTTGAAATTTATATAATAATAATTAATTAATTAATGAACAAATGAA 1188
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 271 TyrLysLeuLysAsp-----LeuLeuArgLeuProSer 282
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1189 ATTTATTATCGAAAAATTTGGGATGTTGATGATATAATTTATTCGTTTAGATAGT 1242
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 123
US-09-662-254B-38
; Sequence 38, Application US/09662254B
; Patent No. 6933145
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Li, Yi
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous
; FILE OF INVENTION: Vertebrate Cells
; FILE REFERENCE: UP-221C1XC1
; CURRENT APPLICATION NUMBER: US/09/662,254B
; CURRENT FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 09/086,651
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/224,479
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Amsacta moorei entomopoxvirus
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(2163)
; OTHER INFORMATION:
US-09-662-254B-38
```

```
Alignment Scores:
Pred. No.: 0.398 Length: 2163
Score: 96.00 Matches: 50
Percent Similarity: 40.4% Conservative: 32
Best Local Similarity: 24.6% Mismatches: 91
Query Match: 6.1% Indels: 30
DB: 3 Gaps: 10

US-10-734-719-9 (1-291) x US-09-662-254B-38 (1-2163)
Qy 92 ValLysThrPheTyrAspTyr-----PheProAspAlaHisLeuGlyTyr 106
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10 ATTAATAATATTTATGATTTTAAACACATATAATTTATTTCCCGAATTCATATAATAT 69
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 107 AspPhePheLysGlnLeu-----LysGluPheAsnAlaTyrPheLys 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 AACTATATTTCCACATTTTATTTCTCTAATAATACTAATAATTTCAATCATATATTGAT 129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 TTTCAT-----TATGTAAAAAATAATAATAATTTTAAATATATTATACCTGTT 183
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 TACAAACTATATTCGAAAAATATGAT-----ATTGTATATAATCAAAATAGTAAT 234
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 161 SerTyrAlaPheAspThrLysGln-----GluAsnLeuLeuLysLeuAlaProAspPhe 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 AAAATATATTTAGATAATAAAGAAATATATAATACAGTATTGAAATTAATTAATGATTTT 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 179 LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGlu 198
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 TTAATAGAT-----GGAATAGATATAATAATAATAATAATACTATATAGA 339
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 199 PheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAla 218
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 TCAACCGGATCTACAAATTTACTTTCTGCATACGCATATGCA-----ACAATATTATAT 393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 219 AsnPheIleGluLeuAlaProAsnLeuAsnSerAsn-----PheIleIleGlnGluLys 236
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 GATTTACCATATAGATTAGGAAATTTAGATATAATCAAAATTTTGGAAATTTGAGAAAGT 453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 237 AsnAsnTyrThrLysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLys 256
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 TCTAATATATTA-----GGAATATTCTACAATGAAGAACAAAAAAGTTTCCTAAA 510
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 257 AsnIleAsnPheLysLysIleLysGluAsnValTyrTyrLysLeuLysAsp 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 511 TATATTAT-----AATATAGATTAGAAAAAATATATTATTAAATTTAAGGAATCT 564
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 277 LeuLeuArg 279
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 AACCTTAGA 573
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 124
US-09-230-380-8
; Sequence 8, Application US/09230380A
; Patent No. 6172215
; GENERAL INFORMATION:
; APPLICANT: Keshi, Hiroyuki
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Nishida, Keigo
; APPLICANT: Matsuhisa, Akio
; TITLE OF INVENTION: Probes For Detecting and Identifying Helicobacter
; FILE REFERENCE: 19036/35268
; CURRENT APPLICATION NUMBER: US/09/230,380A
; CURRENT FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2393
```



```
QY 156 TyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAla 175
Db 838 ---AAACAGATGCTAGTGAATGCTTCTCGTTTACGTGAGAAT----- 797
QY 176 ProAspPheLysAsnAspArgSerHisTyr-----IleGlyHisSerLysAsnThrAsp 193
Db 796 AGTGATTGGAAGATATTCCTATTCTTTGCTATCTATAAACAATCAAGTCAGATTCC 737
QY 194 IleLysAlaLeuGluPheLeu-----GluLysThrTyrLysIleLys 207
Db 736 ATTACACAGGTGAATTTATAGTTGGTACTACGGTTCAGAGGGTAAACATAAATAAC 677
QY 208 LeuTyr-----CysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 676 TCATGGGATAATTAATAAGAAAGACAGCCTTAATCTTCGTCACACTGCAGCTGATTAT 617
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 616 GATGAA-----ACGTTGAATAATAACTTT-----AAACAGTTTAATGATAATTG 572
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 571 CAATCGTATTTTCAAACTTCACACAGCAGTTGGTAAGGT---AAATTCGTAATAAA 515
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLys 275
Db 514 AAAGCTAAACAACACTTACAGTTGATTGCCC-TATAGATTATTACGG 471
RESULT 126
US-09-949-016-12896/c
; Sequence 12896, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12896
; LENGTH: 451924
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12896
Alignment Scores:
Pred. No.: 806 Length: 451924
Score: 96.00 Matches: 70
Percent Similarity: 38.2% Conservative: 35
Best Local Similarity: 25.5% Mismatches: 99
Query Match: 6.1% Indels: 70
DB: 3 Gaps: 17
US-10-734-719-9 (1-291) x US-09-949-016-12896 (1-451924)
QY 48 ValPheTyrThrProAsnPhePhePhe---GluGlnTyrTyrThrLeuLys----- 63
Db 415370 GTGATAAACACATCAAAACATTTTTTTTAAATCAACAACCTGAAACCATCAACTTTAAGAC 415311
QY 64 -----HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyr 81
Db 415310 ATGTTGGTACTGTATCCCTTTCAAAATGATATATAGTTTCATGTTAAATTAAGAACTAT 415251
QY 82 Asn---GlnAlaHisLeuGlu-----AsnGluAsn----- 90
Db 415250 CATATTATTTTACATTTTGAATAATCTGTACATACTAAACAGTTAATTAGTGTATAATA 415191
QY 91 ---PheValLysThr-----PheTyrAspTyrPheProAspAlaHisLeu 104
Db 415190 GAATTTTACAAAACATAAAACTCACTTAAATTTCTTGAATAACTTTTCCA-----GCACGTG 415137
QY 105 GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle 124
Db 415136 ATTTTTCATTTCTTTTAAGCATTTCTATTATTTTATGCG-----AAATA 415092
QY 125 TyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly 144
Db 415091 TATTTAGATGAAAGTCAACCAACTGTTTGTAGTTCTCATATAATAATGAAGCTCAG 415032
QY 145 TyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPhe 164
Db 415031 TTTAAGAAATAT-----AsnPheIleGluLeu-----GATTATCTGTG 415008
QY 165 AspThrLysGlnGlu-----Asn-Leu---LeuLysLeuAlaProAsnPheLysAsnAs 181
Db 415007 GTCACCTTACCAGGAATAGTAACCTTTAGAATACACCTCCTCACAATAATGTCAGAGA 414948
QY 181 pArgSerHisTyrIleGlyHisSerLys-----AsnThrAspIleLysAlaLeuGluPh 199
Db 414947 TAGTCAGTTCTAGATCCCTTTCAAAGCCAAATAAACCTGGGTTTAAAAAAGTAGAA 414888
QY 199 eLeuGluLysThrTyrLysIleLysLeuTyrCys-----LeuCysProAsnSerLeuLe 217
Db 414887 TATTGAAGGGATGTTTAAAAAACAGTTACATTTGTTACAGCCTATCTCCGAATAGCATGAT 414828
QY 217 uAla-----AsnPheIleGluLeu----- 223
Db 414827 GACTCCCCAACCCCACTCAAGCAGGTACTAGAGGTTTATTCTCTGGAGAGGAGCAAGA 414768
QY 224 ---AlaProAsnLeuAsnSer-AsnPheIleIleGlnGluLysAsnAsnTyrThrLysA 242
Db 414767 CGATGCTGGACAGTTGAACAGGAGACCATATATTAAACACAGAGAATAAGTTCAAAT 414708
QY 242 spile---LeuIleProSerSerGluAlaTyr-----GlyLysPheSerLysAsnI 258
Db 414707 CCATAACCACTGTTTGAGACCACTGAGCCTTATGCTGTGTTGATTATCATATCTCTCCA 414648
QY 258 leAsnPheLysLysIleLysIleLysGluAsnValTyr 270
Db 414647 GAAAGATAATAGATAATTTTTTTCAGCAAAATAAATAT 414610
RESULT 127
US-09-949-016-17305/c
; Sequence 17305, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17305
; LENGTH: 451925
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17305
Alignment Scores:
Pred. No.: 806 Length: 451925
Score: 96.00 Matches: 70
Percent Similarity: 38.2% Conservative: 35
Best Local Similarity: 25.5% Mismatches: 99
Query Match: 6.1% Indels: 70
DB: 3 Gaps: 17
US-10-734-719-9 (1-291) x US-09-949-016-12896 (1-451924)
QY 48 ValPheTyrThrProAsnPhePhePhe---GluGlnTyrTyrThrLeuLys----- 63
Db 415370 GTGATAAACACATCAAAACATTTTTTTTAAATCAACAACCTGAAACCATCAACTTTAAGAC 415311
QY 64 -----HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyr 81
Db 415310 ATGTTGGTACTGTATCCCTTTCAAAATGATATATAGTTTCATGTTAAATTAAGAACTAT 415251
QY 82 Asn---GlnAlaHisLeuGlu-----AsnGluAsn----- 90
```

Pred. No.: 806 Length: 451925  
Score: 96.00 Matches: 70  
Percent Similarity: 38.3% Conservative: 35  
Best Local Similarity: 25.5% Mismatches: 99  
Query Match: 6.1% Indels: 70  
DB: 3 Gaps: 17

US-10-734-719-9 (1-291) x US-09-949-016-17305 (1-451925)

```
QY 48 ValPheTyrThrProAsnPhePhePhe---GluGlnTyrTyrThrLeuLys----- 63
Db 415370 GTGATAACACATCAAAACATTTTAAATCACAACCTCAAAACCATCAACTTTAAGAC 415311
QY 64 -----HisLeuIleGlnAsnGlnGluThrGluThrGluLeuIleMetCysSerAsnTyr 81
Db 415310 ATGTTGCTACTGTATCCCTTCAAAATGATATATAGTTTCATGTTAAATTAAGAACTAT 415251
QY 82 Asn---GlnAlaHisLeuGlu-----AsnGluAsn----- 90
Db 415250 CATATTATTTCCATTTGAAATATCTGATCTACATATCAATACAGTTAATTAGTGTAATAA 415191
QY 91 ---PheValLysThr-----PheTyrAspTyrThrProAspAlaHisLeu 104
Db 415190 GAAATTTACAAAACATAAACTCACCTTAAATTTCTTGAATATCTTTCCA-----GCACGTG 415137
QY 105 GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle 124
Db 415136 ATTTTTCATTTCTTTAAGCATTTCTATTATTTTAAATGCG-----AAAATA 415092
QY 125 TyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly 144
Db 415091 TATTTAGATAGAAAGTCAACCAACTGTTTGTAGTTCTCATATTAATAATGAAAGCTCAG 415032
QY 145 TyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPhe 164
Db 415031 TTTAAGAAATAT-----GATTATCTGCTG 415008
QY 165 AspThrLysGlnGlu-----Asn-Leu---LeuLysLeuAlaProAspPheLysAsnAs 181
Db 415007 GTCACCTACCGAGGAATAGTAACCTTTAGAAATACACCCCTCACTCCCAAAATGTCCAGAGA 414948
QY 181 pArgSerHisTyrIleGlyHisSerLys-----AsnThrAspIleLysAlaLeuGluPh 199
Db 414947 TAGTCAGTTCTAGATCCCTTTCAAAGCCAAATAAACGTGGGTTTAAAAAAGTGAAT 414888
QY 199 eLeuGluLysThrTyrLysIleLysLeuTyrCys-----LeuCysProAsnSerLeuLe 217
Db 414887 TATTGAAGGGATGTTTAAAAAACAGTTACATTTGTACAGCCTATCTCCGAATAGCATGAT 414828
QY 217 uAla-----AsnPheIleGluLeu----- 223
Db 414827 GACTCCCCCAACCCACTCAAGCAGGTACTAGAGGTTTATTTCTCGAGAGGAAGCAAGA 414768
QY 224 ---AlaProAsnLeuAsnSer-AsnPheIleIleGlnGluLysAsnTyrThrLysA 242
Db 414767 GCATGCTCGACAGTGAACCGAGGAGACCATATATTAAACCAAGAAATAAGTTGAAT 414708
QY 242 spIle---LeuIleProSerSerGluAlaTyr-----GlyLysPheSerLysAsnI 258
Db 414707 CCATACACCATGTTGAGACCATGAGCCTTATGCTGCTTGATTTCATATACTTCTCTCCA 414648
QY 258 leAsnPheLysLysIleLysIleLysGluAsnValTyr 270
Db 414647 GAAAGAATAATAGATATTTTTCAGCAAAATAAATAAT 414610
```

RESULT 128

US-09-248-796A-4739  
; Sequence 4739, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstock et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 4739  
; LENGTH: 2379  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-4739

Alignment Scores:  
Pred. No.: 0.596 Length: 2379  
Score: 95.00 Matches: 49  
Percent Similarity: 41.4% Conservative: 30  
Best Local Similarity: 25.7% Mismatches: 72  
Query Match: 6.1% Indels: 40  
DB: 3 Gaps: 9

US-10-734-719-9 (1-291) x US-09-248-796A-4739 (1-2379)

```
QY 109 PheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle-----Tyr 125
Db 598 TATGAAAGATGAAAGAA-----GAACACGAAAGAGCCACGAGTTGACTACTTCGGTT 651
QY 126 PheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyr 145
Db 652 TTCTCAAGAAAGAGAACCCCTTAAT-----TCCGAGTCGAAACAATAT 693
QY 146 LysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAsp 165
Db 694 AAAGAGCAATCGCTGAACACAGCTCAATTTGAAGAGAGAAATATTATCCAAAAATGATACT 753
QY 166 ThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyr 185
Db 754 GTCAAAAAGATCAATTTGTATAGCTT-----TTCCATTAATGAAAGAAACACAAT 804
QY 186 -----IleGlyHisSerLysAsnThrAspIle-----LysAlaLeu 197
Db 805 CAATTAAGAGATGAAATTAATCAAAAACGAGGAGTTGAAACTTGCTAAGAAAGAGTTA 864
QY 198 GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu 217
Db 865 TCGAATAAGAGAAACATACAAA-----TCCATAATGAAAGAAACACAAT 897
QY 218 AlaAsnPheIleGluLeuAlaProAsnLeu-----AsnSerAsnPheIleIle 233
Db 898 ACAGATTACTCAAGCAGTGTGTTGAATGCCAAAAGCAAAAGCAAAACAGAGGACGT 957
QY 234 GlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLys 253
Db 958 CAACAGAAATCGACTCAACGAAACCGTAATTAATCCAGTTGAGCTAACCCAGAGATCA 1017
QY 254 PheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeu 273
Db 1018 TTATCCAAATAAAATCAATTTCCATAAACTTAAGGTTGGTGAC-----TTG 1062
QY 274 IleLysAspLeuLeuArgLeuProSerAspIle 284
Db 1063 GTAAAGGATATTGCTAGACAAAGACGCAAGTT 1095
```

RESULT 129

US-09-710-279-4137/c  
; Sequence 4137, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279

```

; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4137
; LENGTH: 3131
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
;
US-09-710-279-4137

Alignment Scores:
Pred. No.: 0.882 Length: 3131
Score: 95.00 Matches: 73
Percent Similarity: 37.1% Conservative: 43
Best Local Similarity: 23.3% Mismatches: 107
Query Match: 6.1% Indels: 91
DB: 3 Gaps: 16

US-10-734-719-9 (1-291) x US-09-710-279-4137 (1-3131)

Qy 31 AsnGlnPheTyrPheGluAspLys-----TyrTyr----- 40
Db 2831 AATCAATTAAATTTCTAATCAAGATTAAATCACTACTACTACAAATATATCGTCAACTAGT 2772
Qy 41 -----LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePhe 56
Db 2771 ATCGGAATTAAATGTTTCAAAGCCAAATAAATATTTTTTTTGACATAATCATCTTTTACC 2712
Qy 57 GluGlnTyr-----TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGlu 72
Db 2711 TAAAAATTTACAAAATTAGCTACACATTTAAACAAAATGCTAAAGAGGTTATATG 2652
Qy 73 -----ThrGluLeuIleMetCysSerAsnTyr 81
Db 2651 TGAAGAAAGTTGTCGATAATAACATATTAATAATAAGAGGAATAATATCTTCTGAAATAC 2592
Qy 82 AsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyr---AspTyrPhePro 100
Db 2591 AATCAA-----TAGACTTACTACAGAGATTACACAGC 2559
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 2558 TT-TTAGTTCAGGATTCCTTATATCAAGATTTTAAATTTTAACTCCTATTTTAAA 2500
Qy 121 PheHisGluIleTyrPheAsn-----GlnArgIleThrSerGlyValTyrMet 136
Db 2499 TATAAGAGCGCAACAATAATAATAAAGATTATCCAACATCTTACAAAACGGTGTACATGT 2440
Qy 137 CysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr 156
Db 2439 TATGATATTTTAAAAATAAATAGGATTCACGAATTAGTTCCTTCAAAATAAAATTTGCT 2380
Qy 157 GlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaPro 176
Db 2379 GAAAAC-----TACGGCAACATTGAGGAGGCTCTCGTT----- 2347
Qy 177 AspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAla 196
Db 2346 -----GATACTGTTCAATATATG-----AAAAAGAAATCTGAAAGCT 2311
Qy 197 LeuGluPheLeuGluLysTyrThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu 216
Db 2310 AAAAAACGACTCATCAAAACCTTACAAT-----CCTGTTTGCATTA 2265
Qy 217 LeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys 236
Db 2268 ATTTCTATCTTCTTATTCATATTAACCAATTTTAAATATAACTGTCATACCTCAATTTCAA 2209
Qy 237 AsnAsnTyrThr-----LysAspIleLeu----- 244

```



```
Db 937 TATAAGAGCGAACAATAAATAAAGAGATTATCCAAACATCTACAAAACGGTGCTACATGT 996
Qy 137 CysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr 156
Db 997 TATGATATTTAAATAATAATAGGGTATCCAGAAATTAGTTCTTCTTCAATAAATTTTCT 1056
Qy 157 GlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaPro 176
Db 1057 GAAAAC-----TAGCGCAACATTGAGGAGGCTCTCGTT----- 1089
Qy 177 AspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAla 196
Db 1090 -----GATACTGTTCAATATATG-----AAAAGAAATCTGAAGCT 1125
Qy 197 LeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu 216
Db 1126 AAAAAGACGACTCATCAAAACCTTCAATAT-----CCTGTTGCATTA 1167
Qy 217 LeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys 236
Db 1168 ATTTCATCTCTTATTTCATATTAACCATTTTAAATATATACTGTCATACCTCAATTTCAA 1227
Qy 237 AsnAsnTyrThr-----LysAspIleLeu----- 244
Db 1228 CAACTTATGAGCATATGAATGTTAAATATCAACATTTCAAAATCTACTACTTATT 1287
Qy 245 -----IleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLys 261
Db 1288 ATTACCGCTCTTCCAAACTAACTTTCATTTTATCTTTTATAGTGGTATAGCATTT--- 1344
Qy 262 LysIleLysIleLysGluAsnValTyrTyr-----LysLeuIleLysAsp 276
Db 1345 -----TTTATCACTATATAAATCTCTACTATTATCTACCAATTTGAGAAAAAGTTAAATCT 1398
Qy 277 LeuLeuArgLeuProSerAspIleLysHisTyrPheLys 289
Db 1399 ATTTTAAAAATCCCAATG---ATTAATAGCTATTATAAA 1434
```

## RESULT 131

```
US-09-710-279-2185
; Sequence 2185, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2185
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-2185
```

```
Alignment Scores:
Pred. No.: 0.181 Length: 939
Score: 94.50 Matches: 57
Percent Similarity: 35.8% Conservative: 43
Best Local Similarity: 20.4% Mismatches: 104
Query Match: 6.0% Indels: 75
DB: 3 Gaps: 13
```

US-10-734-719-9 (1-291) x US-09-710-279-2185 (1-939)

Qy 42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrThr 61

```
Db 55 GCAGAAGAGTCACACCAACGTTTTCGATGGGTCA-----AGCCAACAACCTTTTGTCA 108
Qy 62 LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyr 81
Db 109 TTATATCATATGTTTACCTTAACCTTAAAGTATTCAACT-----TTT 147
Qy 82 AsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPhe----- 95
Db 148 GAGCAATTATATCTCAAAAGGATGATTTTAAAAAATCAATTCAAGTAAATAATACATCCA 207
Qy 96 -----TyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeu 112
Db 208 CAATACACTTATGATAGTCTCACTCAAACTTTAGTTGCATACAAATTTACTTATCCAAACG 267
Qy 113 -----LysGluPheAsnAlaTyrPheLysPhe---HisGluIleTyrPhe 126
Db 268 TTATCTCATACACGCAAGGAGTCAAAATACATATTATTCCAATCGTTCAAAATACCTATATC 327
Qy 127 AsnGlnArgIleThrSerGlyValTyrMetCysAlaVal-----AlaIleAlaLeu 143
Db 328 CAACAAAGGTTTAAACAA---CTTTATCATCAGGTCATTTGAATCAATCAAGTATCAAT 384
Qy 144 GlyTyrLysGluIleTyrLeu-----SerGly 152
Db 385 ACTATAGACGAAATATATTATTATTGAGAATTTTAAATAATAATAATACCATACATTT 444
Qy 153 IleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGlu---AsnLeu 171
Db 445 CTTCAATTATTATTCAAGGATATGAGGAATCCATGTATATCTACTAGACAAACAATAAGTTTA 504
Qy 172 LeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsn 191
Db 505 ATTGAGAGTATACCA-----CAATCAGATATTATTGACGAGAAATGAAT 549
Qy 192 ThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211
Db 550 GNACTGATTGACATATTGATCAATTAAGAGATTCACGGAATAT----- 594
Qy 212 CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231
Db 595 -----CCAATACTTTCTCAAGCTATCAATCTTTCCACCATTTACTTAAACAATACATAC 645
Qy 232 IleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr 251
Db 646 TTA-----AGCTAT 654
Qy 252 GlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyr 271
Db 655 CAAAAGTTAAATCTGGTCTCAATTTAAAGAAATTTGCTCAATTACAAAATGTAAACTT 714
Qy 272 LysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyrPheLysGly 290
Db 715 AACACAATTGAAGATCATATT-----CTAGAAATGTATTATAAGGT 756
```

## RESULT 132

```
US-09-134-001C-1054
; Sequence 1054, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1054
; LENGTH: 1950
```

TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-1054

Alignment Scores: 0.514 Length: 1950  
Pred. No.: 94.50 Matches: 51  
Score: 94.50  
Percent Similarity: 37.6% Conservative: 34  
Best Local Similarity: 22.6% Mismatches: 90  
Query Match: 6.0% Indels: 51  
DB: 3 Gaps: 9

US-10-734-719-9 (1-291) x US-09-134-001C-1054 (1-1950)

```
QY 6 IleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAspPhe 25
Db 337 CTTCACGAAATGGACAGATTATAATCTGATCAATCAATCAATCAATCAATCAATCAAT 396
QY 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCys 45
Db 397 GAAACCTTTA--ACGAATCGTTATGAGCAATTAGATGGTTTCAATATGAAAGTAAAT 453
QY 46 LysAlaValPheTyrThrProAsnPhe----- 54
Db 454 AAACAGTCTCTGCACGGACTTAACCTTTACAGAAAGCGATTTTAAACCAACCAATTAATGAT 513
QY 55 PhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGlnTyrGluThrGlu 74
Db 514 TTATGTTGGTGGTCAAAAACACGCTCTTCTCTGCTCAATGTTACTTTAAAGAACCGAT 573
QY 75 LeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThr 94
Db 574 TTACTACTTCTTGATGAA----- 591
QY 95 PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu 114
Db 592 -----CCAACCTAACCATCTAGACATGGAAACCAACCAAGTGGTTAGAGGAT 636
QY 115 PheAsnAlaTyrPheLys-----PheHisGluIleTyrPheAsnGln 128
Db 637 TACTTAAAGTACTTTAAAGGGGCAATTGTAATTAATCATGTCATGATAGATATTTTATAGAT 696
QY 129 ArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148
Db 697 AAAATTGCTACTCAAGTATAT-----GATGTTGCTCTAGGAAATGTAAACAC 744
QY 149 TyrLeuSerGlyIleAsp---PheTyrGlnAsnGlySerSerTyr----- 162
Db 745 TATGTAGGAAACTATGAGCAATTTTATAAGCAACGTAATCACTATTAACGAAACGCGATG 804
QY 163 ---AlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAsp 181
Db 805 CAAGAATTTGAAAGCAACAGAGAAATCAAGCGTTTGAACCTTTGTAGAAAAAAT 864
QY 182 -----ArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 865 ATACTCGTGCATCTACAGTGGCATGCTGCTAAGAGTAGACGT---AAACATTTGAAAAA 921
QY 200 LeuGluLysThrTyrLys 205
Db 922 ATGGAACGATTGATAAA 939
```

## RESULT 133

US-09-710-279-3750  
Sequence 3750, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMMERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P034800S  
CURRENT APPLICATION NUMBER: US/09/710,279  
CURRENT FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3750

LENGTH: 3068  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-3750

## Alignment Scores:

Pred. No.: 0.98 Length: 3068  
Score: 94.50 Matches: 57  
Percent Similarity: 35.8% Conservative: 43  
Best Local Similarity: 20.4% Mismatches: 104  
Query Match: 6.0% Indels: 75  
DB: 3 Gaps: 13

US-10-734-719-9 (1-291) x US-09-710-279-3750 (1-3068)

```
QY 42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr 61
Db 191 GGCAAGAAGTCACACCAACCGTTTTTCGATCGGTCA-----AGCCAACAACCTTTGTCA 244
QY 62 LeuLysHisLeuIleGlnAsnGlnGlnTyrGluThrGluLeuIleMetCysSerAsnTyr 81
Db 245 TTATATCATAGTTTACCTTAACCTTAAAGTATTTCACT-----TTT 283
QY 82 AsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPhe----- 95
Db 284 GAGCAATTTTATACCTCCAAAAGAGTATTTTAAANAATCAATCAAGTAAATAATACATCCA 343
QY 96 -----TyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeu 112
Db 344 CAATACACTTATGATGTCCTCACTCAAACTTTAGTTCATACATACTTACTTATCCAAACG 403
QY 113 -----LysGluPheAsnAlaTyrPheLysPhe-----HisGluIleTyrPhe 126
Db 404 TTATCTCATACACGACGAGGAGTCAAATACATTTATTTCCAAATCGTTCAAAATACCTATATC 463
QY 127 AsnGlnArgIleThrSerGlyValTyrMetCysAlaVal-----AlaIleAlaLeu 143
Db 464 CAACAAAGGGTTAAACAA-----CTTTATCATCAGGTCATTCGAATCAATCAAGTATCAAAT 520
QY 144 GlyTyrLysGluIleTyrLeu-----SerGly 152
Db 521 ACTATAGACGAAATATATTTATTTTGAAGATTTTAAATAATAATAACCATACATATTT 580
QY 153 IleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGlu-----AsnLeu 171
Db 581 CTTCTATTATTATTACAAGGATATGAGGAATCCATGTATACATAGACACAAATAAGTTTA 640
QY 172 LeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsn 191
Db 641 ATTGAGAGTATACCA-----CAATCAGAATTTTGAACGAGAAATGAAT 685
QY 192 ThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211
Db 686 GAACGTGATTGACATATTTGAATCAATTTAAAGAGTTCAACGAAATAT----- 730
QY 212 CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231
Db 731 -----CCAATACCTTCTCAAGCTATCATTTCTTCCACCATTTACTAACAATAATACATAC 781
QY 232 IleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr 251
Db 782 TTA-----AGCTAT 790
QY 252 GlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyr 271
Db 791 CAAAAGTTTAAATCTGGTCTCAATTTTAAAGAAATTTGCTCAATTACAAATGTAAAACTT 850
```



```
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/954,441
/ FILING DATE: 20-OCT-1997
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/323,170
/ FILING DATE: 13-OCT-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/010,409
/ FILING DATE: 29-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Einhorn, Gregory P.
/ REGISTRATION NUMBER: 38,440
/ REFERENCE/DOCKET NUMBER: 015280-113110US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 9636 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 149..9556
/ US-08-954-441-1

Alignment Scores:
Pred. No.: 5,01 Length: 9636
Score: 94.50 Matches: 74
Percent Similarity: 33.3% Conservative: 31
Best Local Similarity: 23.5% Mismatches: 97
Query Match: 6.0% Indels: 114
DB: 3 Gaps: 14

US-10-734-719-9 (1-291) x US-08-954-441-1 (1-9636)

Qy 30 CysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPhe 49
Db 7690 TGTGATTACTTTTGATAAGAAATTTTGTACTTATATAGATCC----- 7734
Qy 50 TyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuGlnAsnGln 69
Db 7735 TACAAAAAACAATTTTCTTTGAAGATAT----- 7764
Qy 70 GluTyrGluThrGluLeuIleMetCys-SerAsnTyrAsnGlnAlaHisLeuGluAsnGln 89
Db 7765 ----ATATATACAAAGTTTTTATGGTGTCT-----CTTGATCACT 7803
Qy 89 uAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAsp----- 107
Db 7804 TAATCAATAAAAAAATACATGAAGAATGGATGATGATCAATTTATTTTATCCTCTCA 7863
Qy 107 ----- 107
Db 7864 TAATGTTATTCATAATGTTGTTACTTAATAATATAGTCACTTATCATCTGCATTAGA 7923
Qy 108 -----PhePheLysGlnLeuLysGluPhe-----AsnAlaTyrPheLysPheHi 122
Db 7924 AGGAGTCTTATTAT-GAAATCAAAAGTTTCTGGAGATGAACAGCTACAAAAAACA 7982
Qy 122 sGluIle-----TyrPheAsnGlnArgIleThrSerGlyValty 135
Db 7983 CTACACTACCAACTGATGGTGTATCAAGTATTTTAAATCCACCATATGTAAGGAAGATA 8042
Qy 135 r----- 135
Db 8043 TAACATTTTCATCTTTTGTGGGAATCTACAACAAAAAACAACAAAGAACACAT 8102

136 -MetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSer-GlyIleAsp 155
8103 CTCTTGCACTTATTCATATACATATATATCAACACAGAAATATTTATTCATGATGATT 8162
155 heTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuA 175
8163 TCTTATAT-----TTAGAAAATCAACAAATGATG 8192
175 laProAspPheLysAsnAspArgSerHisTyrIle---GlyHisSerLysAsnThrAspI 194
8193 CTATTAGTAATAATAATAATTCATATCTATATTTTACACATATAATAAATACAGAGA 8252
194 leLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProA 214
8253 ATAATCTAATATGATATATCTTTAATTCACAAAACCTGTTATAGGAATTAATGCTTA 8312
214 snSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleG 234
8313 ATAAAAAATTA-----AATCCACAAACATGTTT----- 8341
234 lnGluLysAsnAsnTyrThrLysAspIleLeuIleProSer-----SerGluA 250
8342 --GATGAAGTCTATTATGTTAAACACAGAGATGTACTTCGAAAACTATAACAGCTGATA 8399
250 laTyrGlyLysPheSerLys-----AsnIleAsnPheLysLysIleLysIle- 265
8400 AATATAATACATTTTAGTAAAGACAAATAGGAATATATATTAATAATGCAATCTATT 8459
266 -----LysGluAsnValTyrTyrLysLeuIle 274
8460 ATAATCCAGATCAAAAGGATTAATACATATATCTATTATA 8500

RESULT 136
US-07-991-867B-25/c
; Sequence 25, Application US/07991867B
; Patent No 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESSES:
; ADDRESS: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C3
; TELECOMMUNICATION INFORMATION:
```





; PRIOR APPLICATION NUMBER: US 08/107,755  
; PRIOR FILING DATE: 1993-08-19  
; PRIOR APPLICATION NUMBER: WO 92/14818  
; PRIOR FILING DATE: 1992-02-12  
; PRIOR APPLICATION NUMBER: US 07/827,685  
; PRIOR FILING DATE: 1992-01-30  
; PRIOR APPLICATION NUMBER: US 07/657,584  
; PRIOR FILING DATE: 1991-02-19  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25  
; LENGTH: 1395  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
US-09-370-861A-25

Alignment Scores:  
Pred. No.: 0.365 Length: 1395  
Score: 94.00 Matches: 48  
Percent Similarity: 38.7% Conservative: 34  
Best Local Similarity: 22.6% Mismatches: 76  
Query Match: 6.0% Indels: 54  
DB: 3 Gaps: 10

US-10-734-719-9 (1-291) x US-09-370-861A-25 (1-1395)

```
QY 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCys 45
DB 627 GATCATTTTACTTGTTCGTAATA-----GATAAACACTTT---GATAAAGATAAA 580
QY 46 LysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
DB 579 AAAGCTGCATAC-----TTTTTTAATAGTAGTGGTTTATATACCGAAGACTTATA 532
QY 66 IleGlnAsnGlnGluTyr-----GluThrGluLeuIleMetCysSer----- 79
DB 531 AAACAAATAAAATAATATATGTTTATTGAATCTGATAGCTATATAAAGCCATAAACAC 472
QY 80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
DB 471 TACAATAGTACTCCCTAACTAATAATGCTTATTATATACATGATGACTATCAGAATAT 412
QY 88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
DB 411 TTAAATGATATATTTAAAAAATGTAATAATTATTACTTTTAAATACCTTTTGAATTACAAT 352
QY 107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
DB 351 GATAGTCCCGATTGGGTATGTTTAATAATAATAATTTTATATATATAGTTTATTCAC 292
QY 117 -----AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal 134
DB 291 ATAAATCTAAATTTGAATTTAAAAAATATATATTCTCTAGTGTATTGTTGATTTA 232
QY 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154
DB 231 TTGGCTAGTAGTTATAGAGTGCATATTTATTTCTAGATATGATATATAGTAGTTGAC 172
QY 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeu 174
DB 171 GAAITTTAAAAAT---ACATTGAAATATTTAATAATAAAATAAAGTTTATGGAATTA 115
QY 175 -----AlaProAspPheLysAsn 180
DB 114 ATAGATATGTATAAAAAAATAGTAATATGATGATGATGATGATGATGATGATGATGAT 55
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThr 192
DB 54 GATTATGATAGTTATATTTGATAAATGAGAAGAATAGT 19
RESULT 140
US-09-662-254B-54
; Sequence 54, Application US/09662254B
```

; Patent No. 6933145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Alison Louise  
; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: UF-221C1XCI  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 54  
; LENGTH: 1395  
; TYPE: DNA  
; ORGANISM: Amsacta moorei entomopoxvirus  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: (1)..(1395)  
; OTHER INFORMATION:  
US-09-662-254B-54

Alignment Scores:  
Pred. No.: 0.365 Length: 1395  
Score: 94.00 Matches: 48  
Percent Similarity: 38.7% Conservative: 34  
Best Local Similarity: 22.6% Mismatches: 76  
Query Match: 6.0% Indels: 54  
DB: 3 Gaps: 10

US-10-734-719-9 (1-291) x US-09-662-254B-54 (1-1395)

```
QY 26 AspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCys 45
DB 769 GATCATTTTACTTGTTCGTAATA-----GATAAACACTTT---GATAAAGATAAA 816
QY 46 LysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu 65
DB 817 AAAGCTGCATAC-----TTTTTTAATAGTAGTGGTTTATATACCGAAGACTTATA 864
QY 66 IleGlnAsnGlnGluTyr-----GluThrGluLeuIleMetCysSer----- 79
DB 865 AAACAAATAAAATAATATATGTTTATTGAATCTGATAGCTATATAAAGCCATAAACAC 924
QY 80 -----AsnTyrAsnGlnAlaHisLeuGlu----- 87
DB 925 TACAATAGTACTCCCTAACTAATAATGCTTATTATATACATGATGACTATCAGAATAT 984
QY 88 ---AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106
DB 985 TTAAATGATATATTTAAAAAATGTAATAATTATTACTTTTAAATACCTTTTGAATTACAAT 1044
QY 107 Asp-----PhePheLysGlnLeuLysGluPheAsn 116
DB 1045 GATAGTCCCGATTGGGTATGTTTAATAATAATAATTTTATATATATAGTTTATTCAC 1104
QY 117 -----AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal 134
DB 1105 ATAAATCTAAATTTGAATTTAAAAAATATATATTCTCTAGTGTATTGTTGATTTT 1164
QY 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154
DB 1165 TTGGCTAGTAGTTATAGAGTGCATTTTATTTCTAGATATGATATATAGTAGTTGAC 1224
QY 155 PheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeu 174
DB 1225 GAAITTTAAAAAT---ACATTGAAATATTTAATAATAAAATAAAGTTTATGGAATTA 1281
QY 175 -----AlaProAspPheLysAsn 180
```

Alignment Scores:	
Pred. No.:	0.78
Score:	94.00
Length:	2378
Matches:	66





```
Db 439 AATATTAAATGATCAAAACGATAAAAGAGAGGCTAATAGCTATTATTAGAGCTATTAAAAGAT 498
QY 277 Leu 277
Db 499 ATT 501

RESULT 144
US-09-830-230A-320
; Sequence 320. Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 320
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-830-230A-320

Alignment Scores:
Pred. No.: 0.13 Length: 615
Score: 93.50 Matches: 56
Percent Similarity: 38.2% Conservative: 51
Best Local Similarity: 20.0% Mismatches: 72
Query Match: 6.0% Indels: 101
DB: 3 Gaps: 15

US-10-734-719-9 (1-291) x US-09-830-230A-320 (1-615)
QY 9 AsnGlyProSerLeuLysGluLeuMetCysSerAsnTyrSerArgLeuProAsnAspPheAspValPhe 28
Db 28 AATTATCCGGAGATAAAAGAGCTTGATTAT----- 57
QY 29 ArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysAlaVal 48
Db 58 AAGATAAATTATTATTCTAGTAAATCGC-----TTA 90
QY 49 PheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLeuGlnAsn 68
Db 91 GATTACTCTATGAGTTT----- 108
QY 69 GlnGluTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsn 88
Db 109 ---GATTTTGCATTAAGATTATAAATCAAAAGATGTTTTTAAATATATCAATAGAGAT 165
QY 89 GluAsn-----PheValLys-----ThrPheTyrAspTyrPheProAsp 101
Db 166 AAGAACACTAATGAGTTTATCAAGTATTATAATAATATATAGCTCTTTTATTTCAT 225
QY 102 AlaHisLeuGlyTyrAsp-----PhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 226 TCTAGCTTCTGAAAGGATATCTATATGTAAGGATTTGAGG---TTTAAATTTTTCAT 282
QY 120 LysPheHisGluLeuTyr-----PheAsnGlnArgIleThrSerGly 133
Db 283 AAAAATTGGAAGATTTTACCTCTGTTCTGTTTGTGTAAGGGCATGAGA----- 336
```

```
QY 134 ValTyrMetCysAlaValAlaAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIle 153
Db 337 GTATACAAATAGAGAGCTTGTTATTCTTTGGGTATGTCAAAATATGATTAGATGATGTT 396
QY 154 AspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLys 173
Db 397 -----CACAAATTATGATATAAGTCTAAAGATATGCAATGTTA--- 435
QY 174 LeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAsp 193
Db 436 -----AACAAAGTTAAAGCAATTCAAA----- 456
QY 194 IleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysPro 213
Db 457 -----GTATTTTGTAAACTTATAAAGACAAACTACAT-----CCG 495
QY 214 AsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIle 233
Db 496 GTCTCTTCAGTTGTTAGAAATTGATTCAATA----- 525
QY 234 GlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLys 253
Db 526 -----GATATTCTA-----GAGATTGATAAGCA 549
QY 254 PheSerLysAsnIleAsnPheLysLysIleLysLysGluAsnValTyrTyrLysLeu 273
Db 550 TTTGATAATTACATAAGTTTTTATTATGTCGAAAAAATTCAAATCTTTTTTTAAAGTT 609

RESULT 145
US-09-830-230A-319
; Sequence 319. Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 319
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-830-230A-319

Alignment Scores:
Pred. No.: 0.142 Length: 654
Score: 93.50 Matches: 56
Percent Similarity: 38.2% Conservative: 51
Best Local Similarity: 20.0% Mismatches: 72
Query Match: 6.0% Indels: 101
DB: 3 Gaps: 15

US-10-734-719-9 (1-291) x US-09-830-230A-319 (1-654)
QY 9 AsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPhe 28
Db 67 AATTATCCGGAGATAAAAGAGCTTGATTAT----- 96
QY 29 ArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysAlaVal 48
Db 97 AAGATAAATTATTATTCTAGTAAATCGC-----TTA 129
```

```
Qy 49 PheTyrThrProAsnPhPheGluInTyTyThrLeuLysHisLeuLeuGlnAsn 68
Db 130 GATTACTCTAGAGTTT-----147
Qy 69 GlnGluTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsn 88
Db 148 --GATTGTCGAATTAAGTTATAAATTCAAAAGATGTTTTTAAATTTATCAATAGAGAAT 204
Qy 89 GluAsn-----PheValLys-----ThrPheTyrAspTyrPheProAsp 101
Db 205 AAGAACCACTAATGATTTTATTCAAGTGATTAATAATAATTATAGCTCTTTTTTTTATGAT 264
Qy 102 AlaHisLeuGlyTyrAsp-----PhePheLysGlnLeuLysGluPheAsnAlaTyrPhe 119
Db 265 TCTACGCTTGGAAGGATTTCTATATTGTAAGGATTTGAGG--TTTAATTTTTTTTATGAT 321
Qy 120 LysPheHisGluIleTyr-----PheAsnGlnArgIleThrSerGly 133
Db 322 AAAACTTTTGAGATTTTACCTCATGTCGTCTTTTGTATGAAAGGCATGAGA-----375
Qy 134 ValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLysSerGlyIle 153
Db 376 GTATACAATAGAGAGCTTGTTATTCTTTGGGTATGTCAAAATATGATTTAGATGATGTT 435
Qy 154 AspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLys 173
Db 436 -----CACAAATTATGATATAAGTCTTAAAGATATGGAATGTTA---474
Qy 174 LeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAsp 193
Db 475 -----AACAAAGTTAAGCAATTCAAA-----495
Qy 194 IleLysAlaLeuGluPheGluLysThrTyrLysIleLysLeuTyrCysLeuCysPro 213
Db 496 -----GTATTTTGTGTTAAACTTATAAAGACAACTACAT-----CGG 534
Qy 214 AsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIle 233
Db 535 GTCTCTTCAGTTGTTAGAAATTGATTCAATA-----564
Qy 234 GlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLys 253
Db 565 -----GATATTCTTA-----GAGATTGATTAAGCA 588
Qy 254 PheSerLysAsnIleAsnPhelLysIleLysGluAsnValTyrTyrLysLeu 273
Db 589 TTTGATAATTACATAAGTTTTTATTATGTCGAAAAAAATTCAAATCTTTTTTTTAAAGTT 648
RESULT 146
US-09-134-001C-754
; Sequence 754, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 754
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-754
Alignment Scores: 0.295 Length: 1095
Pred. No.:
```

```
Score: 93.50 Matches: 49
Percent Similarity: 39.7% Conservative: 34
Best Local Similarity: 23.4% Mismatches: 75
Query Match: 6.0% Indels: 51
DB: Gaps: 10
US-10-734-719-9 (1-291) x US-09-134-001C-754 (1-1095)
Qy 105 GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle 124
Db 133 GATTCACCTTATATCAAGTTTAAATTTTAACTCTTATTTAAATATAAAGAGCGA 192
Qy 125 TyrPheAsn-----GlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 193 ACAATAATAAAGATTTATCCAACTCTCAAAACGCTGCTACATGTTATGATATTTA 252
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 253 AAAATAATAGGATCCAGAAATTTCTTTCAAATAAAATTTGCTGAAAC-----306
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 307 -----TACGGCAACATTGAGAGGCTCTCGTT-----333
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 334 GATACCTGTTCAATATATG-----AAAAGAAATCTGAAAGCTAAAAAACGACTC 381
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 382 ATCAAAACCTTACAATAT-----CCTGTTGCATTAATTTCTATCTTC 423
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 424 TTATTCATATTAACCAATTTTAAATATAACTGTCATCTCAATTTCAACAACCTTTATGAG 483
Qy 241 -----LysAspIleLeu-----Ile 245
Db 484 ACTATGAATGTTAAATTTATCAACATTTCAAAATCTACTACTCTTATTTATACCGCTCTT 543
Qy 246 ProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysIleLysIle 265
Db 544 CCCAAACTAACTTTCATCTTTTATCTTTATAGTGGTATAGCATTT-----TTTATC 594
Qy 266 LysGluAsnValTyrTyr-----LysLeuIleLysAspLeuLeuArgLeu 280
Db 595 ACTTATAAATTTCTACTATTATCTACCAATTTGAGAAAAAGTTAAAAATCTATTTTAAAAATC 654
Qy 281 ProSerAspIleLysHisTyrPheLys 289
Db 655 CCAATG---ATTAATACGTATTATAAA 678
RESULT 147
US-09-830-230A-100
; Sequence 100, Application US/09830230A
; Patent No. 6902893
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: PCT/US98/12718
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/057,483
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 60/053,344
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/053,377
; PRIOR FILING DATE: 1997-07-22
; PRIOR APPLICATION NUMBER: 60/050,359
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
```



```
QY 64 HisLeuIleGlnAsnGlnIuTyxGluThrGluLeuIleMetCysSerAsnTyrAsnGln 83
Db 331 ATAGTAATAATAGAGATCAGACAGAGAGAAATGCTAGCTAAAGCGAATTTGTTTTT 390
QY 84 AlahisLeuGluAsnGluAsnPheValIysThrPheTyrAsp----- 97
Db 391 GGATATGTAATAATATGAGAAAT-----GGTTTTATGATCTTCCGAATATAATTTT 441
QY 98 -----TyrPheProAsp-----AlahisLeuGly 105
Db 442 GATCTATTTTAAAGAGATATAAATATTTCTCATGCTAGTTTAAGATTAGCTGAATTAATA 501
QY 106 Tyr-----AspPhePheLysGlnLeuLysGluPheAsn 116
Db 502 TATCTTGTAAAGAAATACTGATGCAATTTCTGCATTTAAAGAGATTAAAGAAATTTCT 561
QY 117 Ala----- 117
Db 562 ATCTCAGGTTATGATAGAGATTATGGCTTTTAAAGTAATAAACTTGGAGTAAGTCAT 621
QY 118 -----TyrPheLysPheHis 122
Db 622 TTAACCTTAGAGTCTTTAGGATTTCTTGACAAACAGCGTTTTTGATACATTTGCTCTTAAT 681
QY 123 GluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMet----- 136
Db 682 GACAATATATTGTAACATAATATATGCGGAGCGCTTTTAAAGATATAATTAATAAAAAAT 741
QY 137 ---CysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIle----- 153
Db 742 GATTGTAGAGTCTATCTTAAG---GATAAAAAAGCAATTTTTTAAATGGCAATAGGCGT 798
QY 154 -----AspPheTyrGlnAsnGlySerSerTyrAlaPhe----- 164
Db 799 TTTCGGGATTATAATGGAACAATTTATATGCTGTGTAATAAATGTTGTTTATTATATAGAT 858
QY 165 -----AspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 859 GATGTTGATGGGATTTAAAGCAATAAATGTCGCCGTAATGCT---GATTTTAGCAAT 915
QY 181 -----AspArgSerHisTyrIleGlyHisSerLysAsnThrAsp 193
Db 916 GTACAAGTTTTCCTGCTGTTAAAAATGGAATATTTGTTGGC-----ACTCTA 963
QY 194 IleLysAlaLeuGluPheLeuGlu---LysThrTyrLys-----Ile 206
Db 964 AATCTCGATTATGCTTTTATGATTTAAATAAATTCGAAAAATATACCGCTTGATCTAAT 1023
QY 207 LysLeuTyrCysLeuCys-----ProAsnSerLeuLeuAlaAsnPheIleGlu 222
Db 1024 AAAATTTCTTCACCTCTGCTTTGATAGTTTAAAAAATTTATTATTAGTTGGACAGTTGAC 1083
QY 223 LeuAla-----ProAsnLeu----- 227
Db 1084 AAGGCTATTATAGTTTAATGTCATAAATTTGAAAAAGATTGAACATTTGGATTTTTTT 1143
QY 228 -----AsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 1144 AGCAAAATGATAATGAAAAAATAATTAATTTTATAAAGAATAAAGATAGTTAT--- 1200
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsn--- 259
Db 1201 -----TTTGTGGACATATGCTGGGGGCTTTTGAATTTAAATTTAAATAAA 1248
QY 260 -----PheLysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277
Db 1249 AATAGTTTACAAAAAGCAGCTTATTGCCAATAATATTGATGTTAATTTATTATGATATG 1308
RESULT 149
US-09-830-230A-591
; Sequence 591, Application US/09830230A
; Patent No. 6902893
```

GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Lyme Disease Vaccines  
; FILE REFERENCE: PB481US  
; CURRENT APPLICATION NUMBER: US/09/830,230A  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: PCT/US98/12718  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/057,483  
; PRIOR FILING DATE: 1997-09-03  
; PRIOR APPLICATION NUMBER: 60/053,344  
; PRIOR FILING DATE: 1997-07-22  
; PRIOR APPLICATION NUMBER: 60/053,377  
; PRIOR FILING DATE: 1997-07-22  
; PRIOR APPLICATION NUMBER: 60/050,359  
; PRIOR FILING DATE: 1997-06-20  
; NUMBER OF SEQ ID NOS: 756  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 591  
; LENGTH: 2154  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-830-230A-591

Alignment Scores:  
Pred. No.: 0.775 Length: 2154  
Score: 93.50 Matches: 68  
Percent Similarity: 35.1% Conservative: 43  
Best Local Similarity: 21.5% Mismatches: 94  
Query Match: 6.0% Indels: 111  
DB: 13 Gaps: 13

US-10-734-719-9 (1-291) x US-09-830-230A-591 (1-2154)

QY 75 LeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValIysThr 94  
Db 55 TTAAGTTTATGTTCTTCTGCTTTCGCAAAAAAGAAATCTCAGGCAATAATTTATTAAAGCG 114  
QY 95 PheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu 114  
Db 115 CATTCAAAAAGAGTTGATTAAATAATTTAAATTTGG-----TTATGGAAT 159  
QY 115 PheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyVal 134  
Db 160 TTGATTTATACAAAAAATAATTTTGAAGCATTTTAAACATAGATCAAGTTCTTACATA 219  
QY 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAsp 154  
Db 220 TATGTTGCTTATTATTATTTAAAAAATAGGATTTTGAAGAGAAATTT-----GTAGAG 270  
QY 155 PheTyrGln-----AsnGlySerSerTyrAla----- 163  
Db 271 TATATGAAAAAGCCCATAGCTAATGAGATAGCATTTGCCAGTTTCTCCAGTTTGGGATTAAAG 330  
QY 164 -----PheAspThrLysGlnGluAsnLeu-----Lys 173  
Db 331 CTATTGTAATTTTAACTCAGCAAAAGAGATATTTTGCATCTGCAATTCATTTGGAGAGAG 390  
QY 174 LeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsn----- 191  
Db 391 CTTTATAAAAAATACGAAAAATAAATAATTTATTATCTGGGTACTTTTAAAAAGTCTTTAT 450  
QY 192 -----ThrAspIleLysAlaLeuGluPheLysGluLysThrTyrLysIleLysLeu 208  
Db 451 TGGCAAAACAAAAACGATAAGGCACCTTAGTCTTTTAAATAAGCTTCGATAAGATGAAATTT 510  
QY 209 -----TyrCysLeuCysProAsnSerLeu----- 217  
Db 511 TCTGATTATCAGAAAAATGAAATAATTTTATTTAAAAAGCAGTTCTTTACCTTAATCTTTCT 570  
QY 218 -----AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn--- 230  
Db 571 AATGTAAGTAGTCAAAAAATTTATTTTAAATGACCTTTTTCAGAACCTTTACCTGCAATTTAT 630

QY 231 -----PheilleGlnGluLysAsn----- 237  
Db 631 TTACATGTAAGAGCTTATGATTATTTATTTGAAAATAAGCTAGTATTTTGGTGCA 690  
QY 238 AsnTyrThrLysAspIleuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsn 257  
Db 691 AATTTTAAATCTTTGTTAGATTAAAGTATGAAGTGGCAATGGCAATTTTAAATGGTGCA 750  
QY 258 IleAsn----- 259  
Db 751 ATAAATATATTAAATAAAAAATGTTTAAATGATTATATGACATAACATTGTTAAGT 810  
QY 260 -----PheLys 261  
Db 811 GATGTTTAAAGGCTTTTATTAGTCTGCGCAAGTTTCAAAATGCTTAAACATTTTATTAGT 870  
QY 262 LysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu----- 277  
Db 871 AAAATAAGAGCAATATATAAAATTTATTATTAGTATTCTTAACCTTAGAGAGAAAAAT 930  
QY 278 -----LeuArgLeuProSerAspIleLysHisTyrPheLysGly 290  
Db 931 AATTAGGACTTCTCTT-----TTAAAGAATATCTTGAAGT 969  
RESULT 150  
US-08-480-604A-9  
; Sequence 9, Application US/08480604A  
; Patent No. 5736139  
; GENERAL INFORMATION:  
; APPLICANT: KINK, JOHN A.  
; APPLICANT: THALLEY, BRUCE S.  
; APPLICANT: PADHYE, NISHA V.  
; APPLICANT: FIRCA, JOSEPH R.  
; APPLICANT: STAFFORD, DOUGLAS C.  
; TITLE OF INVENTION: VACCINE AND ANTI-TOXIN FOR TREATMENT AND  
; TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MEDLEN & CARROLL, LLP  
; STREET: 220 MONTGOMERY STREET, SUITE 2200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,604A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/422,711  
; FILING DATE: 14-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/405,496  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/329,154  
; FILING DATE: 25-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/161,907  
; FILING DATE: 02-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/985,321  
; FILING DATE: 04-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/429,791  
; FILING DATE: 31-OCT-1989  
; ATTORNEY/AGENT INFORMATION:

; NAME: INGOLIA, DIANE E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: OPHD-01763  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7101 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..7098  
; US-08-480-604A-9  
Alignment Scores:  
Pred. No.: 4.24 Length: 7101  
Score: 93.50 Matches: 72  
Percent Similarity: 33.9% Conservative: 41  
Best Local Similarity: 21.6% Mismatches: 119  
Query Match: 6.0% Indels: 102  
DB: 2 Gaps: 12  
US-10-734-719-9 (1-291) x US-08-480-604A-9 (1-7101)  
QY 12 SerLeuLysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsn 31  
Db 1552 AGCTTATGTCATTTGACGATGCAAGCTCAAAAGCTCAATTTGAAGAATATAAAGGAAT 1611  
QY 32 GlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThr 51  
Db 1612 -----TATTTTGAAGTTCT-----CTTGGTGAAGATGATACTTGGATTTT---TCT 1656  
QY 52 ProAsnPhePhePheGluGlnTyrTyrThrLeu-LysHisLeuIleGlnAsnGlnGluTyr 71  
Db 1657 CAAATATAGTAGTTGACAGAGAGATCTTTTAGAAAAAATATCTTCATTAGCAAGAGT 1716  
QY 71 rGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPh 91  
Db 1717 TCAGAGAGAGGATATA-----TACACTATATTGTCAG 1732  
QY 91 eValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGly 111  
Db 1733 -----TACACTATATTGTCAG 1749  
QY 111 nLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleTh 131  
Db 1750 TTACAAGGAGATATAAATTTAGTTATGAGCAGCATGTAACCTTATTTGCAAGAGCTCCTTAT 1809  
QY 131 rSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSe 151  
Db 1810 -GATAGTGTA-----CTGTTTCAGAAAAATATAGAAGATTC 1844  
QY 151 rGlyIleAspPheTyrGlnAsn---GlySerSerTyrAlaPheAspThrLysGlnGluAs 170  
Db 1845 AGAAATTCATATTATTATTAATCTCGAGATGCTGGAATACAGAAATAGACAGATATAA 1904  
QY 170 nLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLy 190  
Db 1905 AATTCCTAAGTATAATTTCTGTATAGACCTTAAGATTAATAAATTAACATTTATTGGTCATGTA 1964  
QY 190 s-----AsnThrAspIle----- 194  
Db 1965 AGATGAATTTAATACTGATATATTTTGCAGGTTTGTAGTAGATTATTATCCACAGAAAT 2024  
QY 195 -----LysAlaLeuGluPhe----- 199  
Db 2025 AGAAGCAGCAATAGATTTTAGCTAAAGAGGATATTTCTCCTTAAGTCAATAGAAATAAATTT 2084  
QY 200 -----LeuGluLysThrTyrLysIleLysLe 208

```
Db 2085 ATTAGGATGTAATATGTTTAGCTACTCTATCAACGTAGAGGAGACTTATCCTGGAAAATT 2144
Qy 208 uTyrcysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAs 228
Db 2145 ATTACTTAAAGTTAAAGATAAAATA-----TCAGAAATTAAATGCCCATCTATAAG 2192
Qy 228 nSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIle-LeuIleProSers 248
Db 2193 TCAAGACTCTATTATAGTAAGTGCAAAATCAATATGAAGTTAGATAAATAGTGAAGGAAG 2252
Qy 248 erGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLys-----I 263
Db 2253 AAGAGAATTATTGGATCATTCTGGTGAATGGATAAAATAAAGAAAGAAAGTATTATAAGGA 2312
Qy 263 leLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu-----L 278
Db 2313 TATTTCATCAAAGAATATATATCATTTTAATCCTAAAGAAAATAAAATTACAGTAAATC 2372
Qy 278 euArgLeuProSerAspIleLysHisTyrPheLys 289
Db 2373 TAAAAATTTACCTGAGCTATCTACATTATTACAAG 2407
```

Search completed: August 21, 2006, 18:25:32  
Job time : 5476 secs

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 21, 2006, 16:02:21 ; Search time 6465 Seconds  
(without alignments)  
4317.564 Million call updates/sec

Title: US-10-734-719-9

Perfect score: 1563

Sequence: 1 MKKVIIAGNGPSLKEIDYSR.....KLKDLRLPSDKHYFKGK 291

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xhl  
-Q=/abs/ABSSWEB.spool/US10734719/runat.21082006.142136.12872/app.query.fasta\_1  
-DB=GenEmbl -QWTF=fastap -SUFFIX=P2N.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=1000  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=150 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss03h  
-USER=US10734719 @CGN 1.1 5548 @runat.21082006.142136.12872 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb env.\*

2: gb pat.\*

3: gb ph.\*

4: gb pl.\*

5: gb pr.\*

6: gb ro.\*

7: gb sts.\*

8: gb sy.\*

9: gb un.\*

10: gb vi.\*

11: gb ov.\*

12: gb htg.\*

13: gb in.\*

14: gb om.\*

15: gb ba.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1563	100.0	876	2 AX934431 Sequence
2	1563	100.0	876	2 BD249794 Campyloba
3	1563	100.0	876	2 AR271703 Sequence

4	1563	100.0	876	2 AR481785 Sequence
5	1563	100.0	876	2 AR527384 Sequence
6	1563	100.0	876	2 AR609663 Sequence
7	1563	100.0	876	2 AR689941 Sequence
8	1563	100.0	876	2 AR691837 Sequence
9	1563	100.0	4749	15 AF305571 Campyloba
10	1563	100.0	6047	15 AY661458 Campyloba
11	1563	100.0	24425	15 AY422197 Campyloba
12	1563	100.0	24437	15 AF167344 Campyloba
13	1559	99.7	873	2 AX934436 Sequence
14	1559	99.7	873	2 AR481787 Sequence
15	1559	99.7	873	2 AR527386 Sequence
16	1559	99.7	873	2 AR609665 Sequence
17	1559	99.7	873	2 AR689943 Sequence
18	1559	99.7	873	2 AR691839 Sequence
19	1559	99.7	12390	15 AF401528 Campyloba
20	1557	99.6	873	2 AX934434 Sequence
21	1557	99.6	873	2 AR481786 Sequence
22	1557	99.6	873	2 AR527385 Sequence
23	1557	99.6	873	2 AR609664 Sequence
24	1557	99.6	873	2 AR689942 Sequence
25	1557	99.6	873	2 AR691838 Sequence
26	1557	99.6	11474	15 AF215659 Campyloba
27	1528	97.8	876	2 AX934427 Sequence
28	1528	97.8	876	2 BD249792 Campyloba
29	1528	97.8	876	2 AR271701 Sequence
30	1528	97.8	876	2 AR481783 Sequence
31	1528	97.8	876	2 AR527382 Sequence
32	1528	97.8	876	2 AR609661 Sequence
33	1528	97.8	876	2 AR689939 Sequence
34	1528	97.8	876	2 AR691835 Sequence
35	1525	97.6	876	2 AX934425 Sequence
36	1525	97.6	876	2 BD249791 Campyloba
37	1525	97.6	876	2 AR271700 Sequence
38	1525	97.6	876	2 AR481782 Sequence
39	1525	97.6	876	2 AR527381 Sequence
40	1525	97.6	876	2 AR609660 Sequence
41	1525	97.6	876	2 AR689938 Sequence
42	1525	97.6	876	2 AR691834 Sequence
43	1525	97.6	6047	15 AF216647 Campyloba
44	1525	97.6	6047	15 AY644679 Campyloba
45	1525	97.6	11442	15 AF400048 Campyloba
46	1525	97.6	11474	2 AX934424 Sequence
47	1525	97.6	11474	2 BD249790 Campyloba
48	1525	97.6	11474	2 AR271699 Sequence
49	1525	97.6	11474	2 AR481781 Sequence
50	1525	97.6	11474	2 AR527380 Sequence
51	1525	97.6	11474	2 AR609659 Sequence
52	1525	97.6	11474	2 AR689937 Sequence
53	1525	97.6	11474	2 AR691833 Sequence
54	1525	97.6	11474	15 AF130984 Campyloba
55	1525	97.6	12576	15 AY297047 Campyloba
56	1521	97.3	1303	15 AY681242 Campyloba
57	1514	96.9	876	2 AX934429 Sequence
58	1514	96.9	876	2 BD249793 Campyloba
59	1514	96.9	876	2 AR271702 Sequence
60	1514	96.9	876	2 AR481784 Sequence
61	1514	96.9	876	2 AR527383 Sequence
62	1514	96.9	876	2 AR609662 Sequence
63	1514	96.9	876	2 AR689940 Sequence
64	1514	96.9	876	2 AR691836 Sequence
65	1511	96.7	11455	15 AY044868 Campyloba
66	1496	95.7	12388	15 AF401529 Campyloba
67	1455.5	93.1	12370	15 AY422196 Campyloba
68	1446	92.5	12403	15 AY854153 Campyloba
69	804	51.4	3071	15 AF195055 Campyloba
70	804	51.4	4668	15 AF257460 Campyloba
71	804	51.4	13484	15 AF400047 Campyloba
72	804	51.4	13484	15 AY044156 Campyloba
73	804	51.4	282183	15 CJ11168X4 Continuation (10 o
74	744.5	47.6	110000	15 CP000957_09
75	744.5	47.6	191996	2 CQ873070 Sequence
76	743	47.5	1293	2 BD134499 Lipopolys



77	743	47.5	1293	2	AR473907 Sequence	150	119	7.6	3360	13	AF095904	AF095904 Toxoplasma
78	743	47.5	1293	2	AR489130 Sequence	c 151	119	7.6	12209	15	AY501976	AY501976 Campyloba
79	743	47.5	11164	15	AF130466 Campyloba	c 152	119	7.6	32215	13	CEM0289	Z82064 Caenorhabdi
80	740.5	47.4	10484	15	AY791515 Campyloba	c 153	119	7.6	34996	13	U87145	U87145 Toxoplasma
c 81	730.5	46.7	960	2	CS222474 Sequence	154	119	7.6	110000	12	PFMAL13_26	Continuation (27 o
c 82	730.5	46.7	13379	2	CQ872953 Sequence	155	119	7.6	110000	15	BA000016_04	Continuation (5 of
c 83	730.5	46.7	110000	15	Continuation (5 of	156	119	7.6	252394	13	AE014833	AE014833 Rickettsi
c 84	730.5	46.7	117576	2	CQ873159 Sequence	c 157	118.5	7.6	10671	15	AE008662	Continuation (8 of
c 85	730.5	46.7	349980	2	CQ873071 Sequence	c 158	118.5	7.6	110000	10	AY653733_07	Continuation (2 of
c 86	721.5	46.2	110000	2	Continuation (4 of	c 159	118.5	7.6	110000	12	PFMAL7P1_01	Continuation (2 of
c 87	721.5	46.2	110000	2	Continuation (4 of	160	118.5	7.6	149662	10	AY077834	AY077834 Sheeppox
c 88	721.5	46.2	110000	2	Continuation (4 of	161	118.5	7.6	149662	10	AY077832	AY077832 Sheeppox
c 89	721.5	46.2	110000	15	L42023_03	162	118.5	7.6	150057	10	AY077833	AY077833 Sheeppox
c 90	692.5	44.3	4277	2	CQ872844 Sequence	163	118	7.5	1539	4	AB038194	AB038194 Malaccotr
c 91	692.5	44.3	10024	15	AE006157 Pasteurel	164	118	7.5	1539	4	AB048827	AB048827 Malaccotr
c 92	681.5	43.6	903	2	CS220684 Sequence	c 165	118	7.5	1539	4	AB083094	AB083094 Malaccotr
93	621	39.7	4977	15	HL11C3	c 166	118	7.5	110000	12	PFMAL13_08	Continuation (9 of
94	152.5	9.8	9515	15	AF486554 Campyloba	c 167	118	7.5	110000	15	AE013218_4	Continuation (5 of
95	152.5	9.8	9561	15	AF486556 Campyloba	168	118	7.5	110000	15	AE017245	AE017245 Mycoplasma
96	152.5	9.8	9568	15	AF486552 Campyloba	169	118	7.5	110000	15	AE017245	AE017245 Mycoplasma
97	135	8.6	178273	12	AC005308 Plasmodiu	c 170	118	7.5	152468	4	AC140023	AF325528 Lumpy ski
c 98	135	8.6	250531	13	AE014845 Plasmodiu	c 171	118	7.5	197731	10	AY484669	AY484669 Rabbitpox
c 99	135	8.6	269619	13	CEY51H4A	c 172	118	7.5	250029	13	AE014820	AE014820 Plasmodiu
c 100	132.5	8.5	271546	13	AE014843	c 173	118	7.5	253305	13	PFMAL3P7	AL034559 Plasmodiu
101	129	8.3	15671	15	AF343914	c 174	117.5	7.5	4068	13	AY582941	AY582941 Litomosol
102	128.5	8.2	110000	4	CR382136_02	c 175	117.5	7.5	86826	13	PFMAL3P5	AL034556 Plasmodiu
c 103	128	8.2	348034	13	CR382400	c 176	117.5	7.5	110000	12	PFMAL7P1_11	Continuation (12 o
c 104	127.5	8.2	2500	2	CS039302	c 177	117.5	7.5	110000	15	BA000021_0	BA000021 Wiggleswo
c 105	127.5	8.2	250195	13	AE014831 Plasmodiu	c 178	117.5	7.5	144250	12	CR354553	CR354553 Danio rer
c 106	127	8.1	83110	12	PFMAL13PB	c 179	117.5	7.5	146570	13	AC117076	AC117076 Dictyoste
c 107	127	8.1	236120	10	AF063866	c 180	117	7.5	52139	13	AC006790	AC006790 Caenorhab
c 108	126	8.1	252394	13	AE014833	c 181	117	7.5	70397	13	AC150256	AC150256 Aedes aeg
c 109	126	8.1	349980	2	AX344554	c 182	117	7.5	167226	11	EX537273	EX537273 Zebraphis
c 110	125.5	8.0	104246	15	CP000013_8	c 183	116.5	7.5	110000	12	PFMAL13_02	Continuation (3 of
c 111	125.5	8.0	110000	15	AE015927_23	c 184	116.5	7.5	110000	12	PFMAL13_19	Continuation (20 o
c 112	125	8.0	49704	4	AF402141	c 185	116.5	7.5	110000	12	PFMAL13_26	Continuation (27 o
c 113	125	8.0	110000	12	PFMAL13_03	c 186	116.5	7.5	110000	12	PFMAL7P1_09	Continuation (10 o
c 114	125	8.0	110000	12	PFMAL8P1_11	c 187	116.5	7.5	149599	10	AY077835	AY077835 Goatpox v
c 115	125	8.0	169794	13	AE004688	c 188	116.5	7.5	149723	10	AY077836	AY077836 Goatpox v
c 116	125	8.0	252650	13	AE014847	c 189	116.5	7.5	171187	5	BS000062	BS000062 Pan trogl
c 117	124.5	8.0	348986	15	EX572093	c 190	116.5	7.5	186854	2	CS106434	CS106434 Plasmodiu
c 118	124	7.9	49843	13	AF288092	c 191	116.5	7.5	253305	13	PFMAL3P7	AL034559 Plasmodiu
c 119	124	7.9	257757	13	AE014837	c 192	116	7.4	1818	15	AF267209	AF267209 Candidatu
c 120	124	7.9	341050	13	PFAL929357	c 193	116	7.4	32179	13	AF047663	AF047663 Caenorhab
c 121	123.5	7.9	131183	4	AF494278	c 194	116	7.4	42090	10	VACRHF	D11079 Vaccinia vi
c 122	123	7.9	6128	2	AX458523	c 195	116	7.4	110000	15	AE017308_2	Continuation (3 of
c 123	123	7.9	6129	2	AX251846	c 196	116	7.4	194711	10	AY243312	AY243312 Vaccinia
c 124	123	7.9	6129	2	AX344232	c 197	116	7.4	236120	10	AF063866	AF063866 Melanoplu
c 125	123	7.9	6129	2	AX348969	c 198	116	7.4	251448	13	AE014819	AE014819 Plasmodiu
c 126	123	7.9	110000	15	AE014075_40	c 199	116	7.4	298664	12	AC006702	AC006702 Caenorhab
c 127	122.5	7.8	174133	4	AF083031	c 200	116	7.4	349402	15	EX842644	EX842644 Mycoplasma
c 128	122	7.8	110000	12	PFMAL13_02	c 201	116	7.4	349418	13	CR382398	CR382398 Plasmodiu
c 129	122	7.8	110000	15	AE009951_04	c 202	115.5	7.4	18166	12	PFMAL8PE	AL929364 Plasmodiu
c 130	122	7.8	277703	13	EX538353	c 203	115.5	7.4	110000	12	PFMAL13_01	Continuation (2 of
c 131	122	7.8	349751	13	PFMAL4P3	c 204	115.5	7.4	110000	12	PFMAL13_10	Continuation (11 o
c 132	121.5	7.8	234081	13	PFMAL4P2	c 205	115.5	7.4	110000	12	PFMAL7P1_03	Continuation (4 of
c 133	121.5	7.8	251762	13	AE014851	c 206	115.5	7.4	110000	15	CP000153_07	Continuation (8 of
c 134	121.5	7.8	310779	12	AC005140	c 207	115.5	7.4	110000	15	CP000153_16	Continuation (17 o
c 135	121	7.7	13267	15	AY800272	c 208	115.5	7.4	110000	15	AE017308_0	AE017308 Mycoplasma
c 136	120.5	7.7	110000	12	PFMAL13_24	c 209	115.5	7.4	110000	15	CP000013_5	Continuation (6 of
c 137	120.5	7.7	110000	15	BA000026_11	c 210	115.5	7.4	112714	4	AC153002	AC153002 Medicago
c 138	120.5	7.7	153477	12	AC006278	c 211	115.5	7.4	250663	13	AE014826	AE014826 Plasmodiu
c 139	120.5	7.7	158548	13	PFMAL3P2	c 212	115.5	7.4	252420	13	AE014841	AE014841 Plasmodiu
c 140	120.5	7.7	250823	13	AE014821	c 213	115	7.4	3244	13	PFGLPH	X74988 P.falciparu
c 141	120.5	7.7	251551	13	AE014844	c 214	115	7.4	110000	12	PFMAL13_22	Continuation (23 o
c 142	120	7.7	34750	13	AY217738	c 215	115	7.4	110000	15	AE013218_3	Continuation (4 of
c 143	120	7.7	189274	10	AF095689	c 216	115	7.4	152177	11	AL935114	AL935114 Zebraphis
c 144	120	7.7	252632	13	AE014818	c 217	115	7.4	158548	13	PFMAL3P2	AL034558 Plasmodiu
c 145	120	7.7	273275	13	AE014828	c 218	115	7.4	205813	12	CR385054	CR385054 Danio rer
c 146	120	7.7	341050	13	PFAL929357	c 219	115	7.4	210549	12	CR848731	CR848731 Danio rer
c 147	119.5	7.6	110000	12	PFMAL13_13	c 220	115	7.4	250022	13	AE014824	AE014824 Plasmodiu
c 148	119.5	7.6	110000	12	PFMAL7P1_10	c 221	115	7.4	281723	13	PFAL929359	AL929359 Plasmodiu
149	119.5	7.6	110000	15	AE015927_12	c 222	114.5	7.3	2225	13	AF395329	AF395329 Tenebrio



C 369	110.5	7.1	11474	2	AR271699	Sequence	442	108.5	6.9	944	15	AF110130
C 370	110.5	7.1	11474	2	AR481781	Sequence	c 443	108.5	6.9	1509	15	VNU89759
C 371	110.5	7.1	11474	2	AR527380	Sequence	c 444	108.5	6.9	11126	15	AE002115
C 372	110.5	7.1	11474	2	AR609659	Sequence	c 445	108.5	6.9	30305	13	AF100305
C 373	110.5	7.1	11474	2	AR689937	Sequence	c 446	108.5	6.9	34966	13	U87145
C 374	110.5	7.1	11474	2	AR691833	Sequence	c 447	108.5	6.9	84162	4	T22J18
C 375	110.5	7.1	11474	15	AF130984	Campyloba	448	108.5	6.9	110099	10	AF270937
C 376	110.5	7.1	12576	15	AY297047	Campyloba	449	108.5	6.9	110099	12	PFMAL8P1_10
C 377	110.5	7.1	32085	4	AC156585	Glonus in	c 450	108.5	6.9	110000	15	AE017308_4
C 378	110.5	7.1	50000	2	AR707080	Sequence	c 451	108.5	6.9	110000	15	AE017197_05
C 379	110.5	7.1	50000	2	AX392733	Sequence	c 452	108.5	6.9	110000	15	AE017245_3
C 380	110.5	7.1	50000	2	AX392733	Sequence	c 453	108.5	6.9	110000	15	CP000013_2
C 381	110.5	7.1	98136	13	DQ158858	Bigelowie	c 454	108.5	6.9	126356	12	AC006876
C 382	110.5	7.1	110000	12	TANN3_17	Continuation (18 o	c 455	108.5	6.9	132597	12	AC162162
C 383	110.5	7.1	110000	13	PFMAL1P2_2	Continuation (3 of	c 456	108.5	6.9	196216	4	AF165818
C 384	110.5	7.1	110000	15	AE000657_05	Continuation (6 of	c 457	108.5	6.9	251448	13	AE014819
C 385	110.5	7.1	139259	13	CEV51A2D	AL021497 Caenorhab	c 458	108.5	6.9	293431	12	PFMAL13P4
C 386	110.5	7.1	184137	5	AP005625	AP005625 Homo sapi	c 459	108.5	6.9	348813	15	EX571865
C 387	110.5	7.1	249943	13	AE014823	AE014823 Plasmodiu	c 460	108.5	6.9	349751	13	PFMAL4P3
C 388	110.5	7.1	250053	13	AE014825	AE014825 Plasmodiu	c 461	108	6.9	1234	4	AF119171
C 389	110	7.0	1269	2	AX741903	Sequence	c 462	108	6.9	2469	2	AR707089
C 390	110	7.0	1269	2	AX741905	Sequence	c 463	108	6.9	2469	2	AX392742
C 391	110	7.0	15421	13	PF00MP1RA	X95275 Plasmodium	c 464	108	6.9	5762	15	SMIGA1PRT
C 392	110	7.0	39984	13	AC114258	AC114258 Dictyoste	c 465	108	6.9	29101	15	AY971355
C 393	110	7.0	10994	11	CR407563	CR407563 Zebrafish	c 466	108	6.9	32405	15	AY102622
C 394	110	7.0	110000	2	AR300198_0	Continuation (14 o	c 467	108	6.9	47577	13	AF396436
C 395	110	7.0	110000	4	CR382135_13	Continuation (26 o	c 468	108	6.9	50000	2	AR707081
C 396	110	7.0	110000	12	PFMAL13_25	Continuation (6 of	c 470	108	6.9	75268	15	AB236337
C 397	110	7.0	110000	15	AE000511_05	Continuation (7 of	c 471	108	6.9	110000	12	PFMAL13_04
C 398	110	7.0	110000	15	AE000783_06	L43967 Mycoplasma	c 472	108	6.9	110000	12	PFMAL13_15
C 399	110	7.0	110000	15	L43967_0	Continuation (18 o	c 473	108	6.9	110000	15	AE000783_02
C 400	110	7.0	110000	15	BA000016_17	Continuation (18 o	c 474	108	6.9	110000	15	BA000021_2
C 401	110	7.0	116696	13	PFMAL3P3	Z98547 Plasmodium	c 475	108	6.9	50000	2	AX392734
C 402	110	7.0	150793	10	AF409137	AF409137 Lumby ski	c 476	108	6.9	115891	12	AC181973
C 403	110	7.0	185887	6	AL691443	AL691443 Mouse DNA	c 477	108	6.9	119544	12	AC147012
C 404	110	7.0	215124	12	CR407565	CR407565 Danio rer	c 478	108	6.9	164399	13	PFMAL3P6
C 405	110	7.0	249995	13	AE014840	AE014840 Plasmodiu	c 479	108	6.9	166742	13	CP000218
C 406	110	7.0	250029	13	AE014838	AE014838 Plasmodiu	c 480	108	6.9	195551	12	AC006281
C 407	109.5	7.0	3343	13	AF138960	AF138960 Neospora	c 481	108	6.9	205719	10	AF438165
C 408	109.5	7.0	3566	10	AF395873	AF395873 Mal de Ri	c 482	108	6.9	252420	13	AE014841
C 409	109.5	7.0	13674	13	U00691	U00691 Dictyosteli	c 483	108	6.9	272698	13	PFMAL4P4
C 410	109.5	7.0	13729	13	U00796	U00796 Dictyosteli	c 484	107.5	6.9	888	2	AR347987
C 411	109.5	7.0	14867	13	AE001398	AE001398 Plasmodiu	c 485	107.5	6.9	11474	15	AF215659
C 412	109.5	7.0	33270	13	AC116921	AC116921 Dictyoste	c 486	107.5	6.9	11858	15	AE002136
C 413	109.5	7.0	47296	13	AF160864	AF160864 Tetrahyme	c 487	107.5	6.9	19209	15	AE002136
C 414	109.5	7.0	110000	12	PFMAL7P1_00	AL844506 Plasmodiu	c 488	107.5	6.9	50000	2	AR274444
C 415	109.5	7.0	110000	13	AC116305_3	Continuation (4 of	c 489	107.5	6.9	50000	2	AR707080
C 416	109.5	7.0	110000	15	CP000102_02	Continuation (3 of	c 490	107.5	6.9	50000	2	AX392733
C 417	109.5	7.0	110000	15	CP000025_09	Continuation (10 o	c 491	107.5	6.9	78153	13	CEY53C10A
C 418	109.5	7.0	196149	12	AC004709	AC004709 Plasmodiu	c 492	107.5	6.9	110000	4	CR382126_23
C 419	109.5	7.0	196216	4	AF165818	AF165818 Guillardi	c 493	107.5	6.9	110000	12	PFMAL13P2_0
C 420	109.5	7.0	317511	15	CJ11168X3	AL139076 Campyloba	c 494	107.5	6.9	110000	12	PFMAL7P1_07
C 421	109.5	7.0	349418	13	CR382398	CR382398 Plasmodiu	c 495	107.5	6.9	110000	15	AE001437_23
C 422	109	7.0	1945	2	CS032822	CS032822 Sequence	c 496	107.5	6.9	110000	15	CE543861_20
C 423	109	7.0	6047	15	AY661458	AY661458 Campyloba	c 497	107.5	6.9	110000	15	AE016826_0
C 424	109	7.0	6668	2	AX251171	AX251171 Sequence	c 498	107.5	6.9	110000	15	AE017199_1
C 425	109	7.0	6668	2	AX344754	AX344754 Sequence	c 499	107.5	6.9	110000	15	AP006841_39
C 426	109	7.0	6668	2	AX346120	AX346120 Sequence	c 500	107.5	6.9	110000	15	BA000016_15
C 427	109	7.0	12593	13	AE001418	AE001418 Plasmodiu	c 501	107.5	6.9	110000	15	BA000023_01
C 428	109	7.0	17429	13	AE001416	AE001416 Plasmodiu	c 502	107.5	6.9	110000	15	CP000013_2
C 429	109	7.0	24425	15	AY422197	AY422197 Campyloba	c 503	107.5	6.9	115521	4	AC146793
C 430	109	7.0	24437	15	AF167344	AF167344 Campyloba	c 504	107.5	6.9	123638	12	AC157779
C 431	109	7.0	45292	13	AC006794	AC006794 Caenorhab	c 505	107.5	6.9	159343	12	AC177240
C 432	109	7.0	110000	12	PFMAL13_11	Continuation (12 o	c 506	107.5	6.9	160621	11	BX005014
C 433	109	7.0	110000	12	TANN4_17	Continuation (18 o	c 507	107.5	6.9	183527	11	AC147908
C 434	109	7.0	110000	15	CP000102_07	Continuation (8 of	c 508	107.5	6.9	253151	13	AE014842
C 435	109	7.0	110000	15	AE017245_4	Continuation (5 of	c 509	107.5	6.9	304639	12	AC171582
C 436	109	7.0	142202	4	AC144474	AC144474 Medicago	c 510	107.5	6.9	330050	13	PFA929355
C 437	109	7.0	250029	13	AE014839	AE014839 Plasmodiu	c 511	107	6.8	1899	5	HSU91641
C 438	109	7.0	343050	13	PFA929353	AL929353 Plasmodiu	c 512	107	6.8	6991	13	PFARHPR
C 439	109	7.0	348174	13	CR382399	CR382399 Plasmodiu	c 513	107	6.8	11455	15	AY044868
C 440	108.5	6.9	804	2	AX110812	AX110812 Sequence	c 514	107	6.8	11455	15	AY044868
C 441	108.5	6.9	944	15	EFA238249	AJ238249 Enterococ						

c 515	107	6.8	57176	12	AC175582	AC175582 Xenopus t	c 588	106	6.8	110000	15	AE015927_20	Continuation (21 o
516	107	6.8	110000	12	PFMAL13_10	Continuation (11 o	589	106	6.8	110000	15	AE016826_3	Continuation (4 of
c 517	107	6.8	110000	12	PFMAL7P1_06	Continuation (7 of	590	106	6.8	110000	15	AE017126_12	Continuation (13 o
518	107	6.8	110000	15	AE017198_09	Continuation (10 o	c 591	106	6.8	110000	15	BA000023_03	Continuation (4 of
c 519	107	6.8	120592	5	AC093509	Homo sapi	c 592	106	6.8	110000	15	BA000023_04	Continuation (5 of
520	107	6.8	125828	12	AC135312	AC135312 Medicago	c 593	106	6.8	124882	5	AC108729	AC108729 Homo sapi
c 521	107	6.8	130660	4	AC144564	AC144564 Medicago	c 594	106	6.8	134671	12	CT027807	CT027807 Danio rer
522	107	6.8	137333	4	AL9335167	AL9335167 Zebrafish	595	106	6.8	157089	4	AY958085	AY958085 Staurastr
523	107	6.8	186814	5	AC093516	AC093516 Homo sapi	596	106	6.8	182504	12	AC019180	AC019180 Homo sapi
524	107	6.8	189800	12	CR848027	CR848027 Danio rer	597	106	6.8	191954	13	AC025726	AC025726 Caenorhab
525	107	6.8	250029	13	AE014838	AE014838 Plasmidii	c 598	106	6.8	200368	5	CNS01DUH	AL133244 BAC seque
526	107	6.8	250029	13	AE014839	AE014839 Plasmidii	c 599	106	6.8	201746	12	AC006895	AC006895 Caenorhab
527	107	6.8	250663	13	AE014839	AE014826 Plasmidii	c 600	106	6.8	246237	13	CEY3986A	AL132948 Caenorhab
c 528	107	6.8	250823	13	AE014821	AE014821 Plasmidii	c 601	106	6.8	253924	13	AE014822	AE014822 Plasmidii
529	107	6.8	253151	13	AE014842	AE014842 Plasmidii	c 602	106	6.8	330050	13	PF9A29355	AL929355 Plasmidii
530	107	6.8	335050	13	PF9A29356	AL929356 Plasmidii	c 603	106	6.8	340552	13	PF9A29354	AL929354 Plasmidii
531	107	6.8	349980	2	AX926714	AX926714 Sequence	c 604	106	6.8	349980	2	AX344560	AX344560 Sequence
532	106.5	6.8	3695	8	PF532679	PF532679 Synthetic	605	105.5	6.7	1125	14	AJ868432	AJ868432 Bos tauru
533	106.5	6.8	4037	13	AB121052	AB121052 Plasmidii	606	105.5	6.7	1375	2	E33327	E33327 Candida tro
534	106.5	6.8	4044	13	AB121054	AB121054 Plasmidii	607	105.5	6.7	1375	4	AB023639	AB023639 Candida t
535	106.5	6.8	4045	13	AB121058	AB121058 Plasmidii	608	105.5	6.7	1836	13	BT001273	BT001273 Drosophil
536	106.5	6.8	4049	13	AB121051	AB121051 Plasmidii	c 609	105.5	6.7	2338	13	AF081950	AF081950 Dictyoste
537	106.5	6.8	4049	13	AB121055	AB121055 Plasmidii	c 610	105.5	6.7	3461	2	CQ774160	CQ774160 Sequence
538	106.5	6.8	4049	13	AB121060	AB121060 Plasmidii	c 611	105.5	6.7	6768	2	I89423	I89423 Sequence 1
539	106.5	6.8	4057	13	AB121053	AB121053 Plasmidii	c 612	105.5	6.7	8457	2	I16670	I16670 Sequence 1
540	106.5	6.8	4060	13	AB121057	AB121057 Plasmidii	c 613	105.5	6.7	8457	2	AR214728	AR214728 Sequence
541	106.5	6.8	4068	13	AB121056	AB121056 Plasmidii	c 614	105.5	6.7	8457	10	AAVSPHERE	M77182 Amsacta ent
542	106.5	6.8	5477	13	PFCAATPAS	X71765 P. falcipar	c 615	105.5	6.7	14529	13	AE001397	AE001397 Plasmidii
543	106.5	6.8	13837	15	AE008579	AE008579 Rickettsi	c 616	105.5	6.7	15872	2	AX251274	AX251274 Sequence
c 544	106.5	6.8	29993	2	DD194608	DD194608 Method an	c 617	105.5	6.7	42112	13	CBRG35L02	AC084559 Caenorhab
c 545	106.5	6.8	35438	4	AX825168	AX825168 Sequence	c 618	105.5	6.7	50000	2	AR707083	AR707083 Sequence
546	106.5	6.8	35438	4	AX347307	AX347307 Penicilli	c 619	105.5	6.7	50000	2	AX392736	AX392736 Sequence
c 547	106.5	6.8	38342	2	AX251503	AX251503 Sequence	c 620	105.5	6.7	85075	12	AC150266	AC150266 Aedes aeg
548	106.5	6.8	38342	2	AX344502	AX344502 Sequence	621	105.5	6.7	92758	15	AE017332_8	Continuation (9 of
c 549	106.5	6.8	39428	13	CEK04B12	Z83232 Caenorhabdi	622	105.5	6.7	98471	4	AC160924	AC160924 Medicago
550	106.5	6.8	110000	10	PF653733_06	Continuation (7 of	c 623	105.5	6.7	110000	2	AR271569_00	AR271569 Sequence
551	106.5	6.8	110000	12	PFMAL7P1_02	Continuation (3 of	c 624	105.5	6.7	110000	2	AR584244_00	AR584244 Sequence
c 552	106.5	6.8	110000	12	PFMAL8P1_10	Continuation (11 o	625	105.5	6.7	110000	12	PFMAL8P1_01	Continuation (2 of
c 553	106.5	6.8	110000	13	PFMAL1P2_1	Continuation (2 of	626	105.5	6.7	110000	15	AE000783_08	Continuation (9 of
c 554	106.5	6.8	110000	13	PFMAL1P2_2	Continuation (3 of	627	105.5	6.7	110000	15	CP000082_14	Continuation (15 o
555	106.5	6.8	110000	15	AE017245_6	Continuation (7 of	628	105.5	6.7	110000	15	CR543861_15	Continuation (16 o
c 556	106.5	6.8	110000	15	BA000026_12	Continuation (13 o	629	105.5	6.7	110000	15	L77117_14	Continuation (15 o
c 557	106.5	6.8	112637	11	BX470076_12	BX470076 Zebrafish	630	105.5	6.7	110000	15	AE009951_14	Continuation (15 o
558	106.5	6.8	113880	13	PFMAL3P4	AL008970 Plasmidii	c 631	105.5	6.7	110000	4	AC155886	AC155886 Medicago
c 559	106.5	6.8	120899	12	AP008181	AP008181 Lotus cor	c 632	105.5	6.7	119696	4	AC155886	AC155886 Medicago
560	106.5	6.8	140590	13	DQ158856	DQ158856 Bigelowie	633	105.5	6.7	122931	12	CT033767	CT033767 Medicago
c 561	106.5	6.8	176518	5	AC023933	AC023933 Homo sapi	634	105.5	6.7	127816	15	AF346500	AF346500 Phototrab
562	106.5	6.8	182871	13	AC117176	AC117176 Dictyoste	c 635	105.5	6.7	131616	2	CQ774045	CQ774045 Sequence
c 563	106.5	6.8	186183	5	AC026827	AC026827 Homo sapi	636	105.5	6.7	131680	10	AY522332	AY522332 Agrotis s
c 564	106.5	6.8	189337	12	CEY17B7	Z82850 Caenorhabdi	c 637	105.5	6.7	134416	3	AF020713	AF020713 Bacteriop
565	106.5	6.8	238923	12	AC146289	AC146289 Danio rer	638	105.5	6.7	135683	3	AP008983	AP008983 Bacteriop
c 566	106.5	6.8	258658	13	AE014832	AE014832 Plasmidii	639	105.5	6.7	201375	15	BSUB0012	Z99115 Bacillus su
567	106.5	6.8	314150	15	CJ11168X1	CR139074 Campyloba	c 640	105.5	6.7	250029	13	AE014816	AE014816 Plasmidii
c 568	106.5	6.8	337203	13	CR382401	CR382401 Plasmidii	c 641	105.5	6.7	254050	13	PF9A29358	AL929358 Plasmidii
569	106	6.8	1032	11	AJ715546	AJ715546 Danio rer	c 642	105.5	6.7	340552	13	PF9A29354	AL929354 Plasmidii
570	106	6.8	1068	11	AJ705098	AJ705098 Danio rer	643	105	6.7	1295	13	AF315286	AF315286 Plasmidii
571	106	6.8	1137	11	AJ871609	AJ871609 Oryzias l	644	105	6.7	1408	4	AB103306	AB103306 Nepenthes
572	106	6.8	1140	11	AJ715548	AJ715548 Tetraodon	645	105	6.7	1728	13	AY099486	AY099486 Plasmidii
573	106	6.8	1295	13	AF315284	AF315284 Plasmidii	646	105	6.7	2361	13	AB189725	AB189725 Plasmidii
574	106	6.8	4152	2	AR707077	AR707077 Sequence	647	105	6.7	4383	13	AF094833	AF094833 Plasmidii
575	106	6.8	4152	2	AX392719	AX392719 Sequence	c 648	105	6.7	4567	13	AC024745	AC024745 Caenorhab
576	106	6.8	6861	4	AB028634	AB028634 Flammulin	c 649	105	6.7	6671	13	PFATP3	X65740 Plasmidii
577	106	6.8	8241	13	AF323343	AF323343 Plasmidii	c 650	105	6.7	7037	2	AX346653	AX346653 Sequence
578	106	6.8	8460	13	CEY52D3	Z99276 Caenorhabdi	651	105	6.7	7622	13	AF112367	AF112367 Plasmidii
c 579	106	6.8	17470	13	AE001402	AE001402 Plasmidii	c 652	105	6.7	8303	2	AX345325	AX345325 Sequence
581	106	6.8	17839	13	AE001384	AE001384 Sequence	653	105	6.7	10029	15	AE006817	AE006817 Sulfolobu
582	106	6.8	50000	2	AR707082	AR707082 Sequence	654	105	6.7	13684	13	AE001403	AE001403 Plasmidii
c 583	106	6.8	50000	10	AY653733_02	Continuation (3 of	c 655	105	6.7	14591	13	AE001406	AE001406 Plasmidii
584	106	6.8	110000	12	PFMAL8P1_09	Continuation (10 o	c 656	105	6.7	15349	13	AE001414	AE001414 Plasmidii
c 585	106	6.8	110000	15	AE001437_34	Continuation (35 o	c 657	105	6.7	26689	4	SPAC17G8	Z69795 S.pombe chr
586	106	6.8	110000	15	CP000102_04	Continuation (5 of	658	105	6.7	34122	13	AF100673	AF100673 Caenorhab
587	106	6.8	110000	15	AE013218_0	AE013218 Buchnera	659	105	6.7	38829	15	AE000787	AE000787 Borrelia
							660	105	6.7	43201	15	AE017226_28	Continuation (29 o

661	105	6.7	100389	15	SS100KBP	Y08256 S.solfatari	c 734	104.5	6.7	180450	13	AE014835	AE014835 Plasmodiu
c 662	105	6.7	100643	12	AP007369	AP007369 Lotus cor	c 735	104.5	6.7	180915	4	CTAJ10592	AJ010592 Guillard
c 663	105	6.7	110000	12	AC005504	AC005504 Plasmodiu	c 736	104.5	6.7	187294	12	CTAJ10592	AE010592 Guillard
c 664	105	6.7	110000	2	AR271569_04	Continuation (5 of	c 737	104.5	6.7	190353	13	AC105293	AC105293 Drosophil
c 665	105	6.7	110000	2	AR584244_04	Continuation (5 of	c 738	104.5	6.7	192929	12	AC005505	AC005505 Plasmodiu
c 666	105	6.7	110000	10	AY653733_08	Continuation (9 of	c 739	104.5	6.7	196490	12	AC005507	AC005507 Plasmodiu
c 667	105	6.7	110000	12	PFMAL8P1_07	Continuation (8 of	c 740	104.5	6.7	250029	13	AE014830	AE014830 Plasmodiu
c 668	105	6.7	110000	15	L77117_09	Continuation (10 of	c 741	104.5	6.7	250351	12	AE014830	AE014830 Plasmodiu
c 669	105	6.7	110000	15	L77117_10	Continuation (11 of	c 742	104.5	6.7	250713	13	AE014850	AE014850 Plasmodiu
c 670	105	6.7	110000	12	AC004157	Continuation (4 of	c 743	104.5	6.7	263335	15	CTJ1168X5	AE0139078 Campyloba
c 671	105	6.7	114226	15	AC004710	AC004710 Plasmodiu	c 744	104.5	6.7	303512	13	AE003467	AE003467 Drosophil
c 672	105	6.7	114226	13	AC149938	AC149938 Strongylo	c 745	104.5	6.7	343050	13	AE002113	AE002113 Ureaplasma
c 673	105	6.7	130796	12	AC180104	AC180104 Plasmodiu	c 746	104.5	6.7	343050	13	PFPA29353	AE002113 Ureaplasma
c 674	105	6.7	133704	12	AC180704	AC180704 Strongylo	c 747	104.5	6.7	349107	15	AE006149	AE006149 Pasteurel
c 675	105	6.7	147656	12	AC176792	AC176792 Strongylo	c 748	104	6.7	349742	15	AE002140	AE002140 Ureaplasma
c 676	105	6.7	169546	12	AC004157	AC004157 Plasmodiu	c 749	104	6.7	349742	15	AE002140	AE002140 Ureaplasma
c 677	105	6.7	180915	4	CTAJ10592	AJ010592 Guillard	c 750	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 678	105	6.7	225405	12	AC149938	AC149938 Strongylo	c 751	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 679	105	6.7	250421	13	AE014849	AE014849 Plasmodiu	c 752	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 680	105	6.7	253001	13	AE014834	AE014834 Plasmodiu	c 753	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 681	105	6.7	293782	12	AC006844	AC006844 Caenorhab	c 754	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 682	105	6.7	327650	15	AC006844	AC006844 Caenorhab	c 755	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 683	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 756	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 684	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 757	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 685	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 758	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 686	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 759	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 687	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 760	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 688	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 761	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 689	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 762	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 690	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 763	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 691	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 764	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 692	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 765	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 693	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 766	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 694	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 767	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 695	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 768	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 696	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 769	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 697	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 770	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 698	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 771	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 699	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 772	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 700	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 773	104	6.7	1116	11	AE005092	AE005092 Prochloro
c 701	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 774	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 702	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 775	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 703	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 776	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 704	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 777	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 705	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 778	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 706	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 779	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 707	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 780	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 708	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 781	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 709	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 782	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 710	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 783	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 711	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 784	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 712	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 785	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 713	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 786	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 714	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 787	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 715	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 788	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 716	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 789	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 717	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 790	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 718	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 791	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 719	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 792	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 720	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 793	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 721	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 794	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 722	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 795	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 723	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 796	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 724	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 797	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 725	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 798	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 726	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 799	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 727	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 800	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 728	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 801	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 729	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 802	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 730	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 803	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 731	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 804	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 732	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 805	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops
c 733	104.5	6.7	1131	5	AE006844	AE006844 Caenorhab	c 806	103.5	6.6	2207	13	MOTARYAB	AE000423 Arabidops

807	103.5	6.6	245802	12	AC006279	Plasmidii	880	102.5	6.6	110000	15	AE014074_15	Continuation (16 o
c 808	103.5	6.6	250743	13	AE014836	Plasmidii	881	102.5	6.6	110000	15	AE014074_16	Continuation (17 o
c 809	103.5	6.6	260929	13	AE014852	Plasmidii	882	102.5	6.6	110000	15	AJ938182_23	Continuation (24 o
c 810	103.5	6.6	29081	12	AC006892	Caenorhab	c 883	102.5	6.6	110000	15	BA000016_08	Continuation (9 of
c 811	103.5	6.6	349980	2	AX344557	Sequence	c 884	102.5	6.6	110000	15	BA000034_02	Continuation (3 of
c 812	103	6.6	1283	4	AF434875	Weldenia	c 885	102.5	6.6	110000	15	CP000003_15	Continuation (16 o
c 813	103	6.6	1334	11	AJ704564	Gallus ga	c 886	102.5	6.6	110000	15	CP000017_15	Continuation (16 o
c 814	103	6.6	1824	2	CS192901	Sequence	c 887	102.5	6.6	110000	15	CP000025_03	Continuation (4 of
c 815	103	6.6	3450	2	AR707104	Sequence	c 888	102.5	6.6	134144	13	DQ158857_01	DQ158857 Bgelowie
c 816	103	6.6	3450	2	AX392757	Sequence	c 889	102.5	6.6	140965	12	AC157669	AC157669 Xenopus t
c 817	103	6.6	3572	10	AF433355	Rice blac	c 890	102.5	6.6	142769	5	AL590043	AL590043 Human DNA
c 818	103	6.6	5728	2	AX345002	Sequence	c 891	102.5	6.6	154541	12	AC157672	AC157672 Xenopus t
c 819	103	6.6	5983	15	AX168957	Myocplasm	c 892	102.5	6.6	155927	12	AC147816	AC147816 Xenopus t
c 820	103	6.6	6627	2	AX346476	Sequence	c 893	102.5	6.6	164748	12	AC141800	AC141800 Apis mell
c 821	103	6.6	12646	4	AM180622	Pichia ac	c 894	102.5	6.6	166259	10	AY689436	AY689436 Deerpox v
c 822	103	6.6	34699	15	AF008210	Buchnera	c 895	102.5	6.6	170560	10	AY689437	AY689437 Deerpox v
c 823	103	6.6	36863	13	U64848	Caenorhabdi	c 896	102.5	6.6	184457	15	AF079317	AF079317 Spingomo
c 824	103	6.6	52836	12	AC174105	Strongylo	c 897	102.5	6.6	187863	12	AC159367	AC159367 Bos tauru
c 825	103	6.6	65691	13	PFMAL3P1	Z98748	c 898	102.5	6.6	209175	12	AC163413	AC163413 Bos tauru
c 826	103	6.6	89004	13	CEY6982	Caenorhabdi	c 899	102.5	6.6	223471	11	CR759874	CR759874 Zebrafish
c 827	103	6.6	110000	4	CR380958_13	Continuation (14 o	c 900	102.5	6.6	254436	13	AE014827	AE014827 Plasmidii
c 828	103	6.6	110000	4	AP008207_259	Continuation (260	c 901	102.5	6.6	257757	13	AE014837	AE014837 Plasmidii
c 829	103	6.6	110000	4	AP008207_260	Continuation (261	c 902	102.5	6.6	349980	2	AX344573	AX344573 Sequence
c 830	103	6.6	110000	10	AY653733_09	Continuation (10 o	c 903	102	6.5	1295	13	AF315280	AF315280 Plasmidii
c 831	103	6.6	110000	12	PFMAL13_15	Continuation (16 o	c 904	102	6.5	1803	15	AF267212	AF267212 Candidatu
c 832	103	6.6	110000	15	AE000511_04	Continuation (5 of	c 905	102	6.5	1833	4	AY936589	AY936589 Phyllanth
c 833	103	6.6	110000	15	AE000783_07	Continuation (8 of	c 906	102	6.5	5407	2	AX346993	AX346993 Sequence
c 834	103	6.6	110000	15	CP000123_02	Continuation (3 of	c 907	102	6.5	6222	2	AX345595	AX345595 Sequence
c 835	103	6.6	110000	15	CP000123_03	Continuation (4 of	c 908	102	6.5	8565	13	AY691326	AY691326 Toxothync
c 836	103	6.6	110000	15	AJ749949_02	Continuation (3 of	c 909	102	6.5	10634	15	AE006313	AE006313 Lactococc
c 837	103	6.6	110000	15	BX571857_18	Continuation (19 o	c 910	102	6.5	10687	3	UEU293027	UEU293027 Uncultu
c 838	103	6.6	110000	15	CP000013_3	Continuation (4 of	c 911	102	6.5	34592	3	AY007505	AY007505 Streptoco
c 839	103	6.6	110000	15	CP000025_10	Continuation (11 o	c 912	102	6.5	38976	13	AF063307	AF063307 Caenorhab
c 840	103	6.6	110000	15	CP000033_11	Continuation (12 o	c 913	102	6.5	61133	12	AC174017	AC174017 Strongylo
c 841	103	6.6	156465	4	AP003234	Oryza sat	c 914	102	6.5	110000	2	BD061520_4	Continuation (5 of
c 842	103	6.6	159577	6	AC140211	Mus muscu	c 915	102	6.5	110000	2	AR409405_4	Continuation (5 of
c 843	103	6.6	172003	4	AP002910	Oryza sat	c 916	102	6.5	110000	12	PFMAL13_12	Continuation (13 o
c 844	103	6.6	200247	6	AC131100	Mus muscu	c 917	102	6.5	110000	15	AE015927_18	Continuation (19 o
c 845	103	6.6	250053	13	AE014825	Plasmidii	c 918	102	6.5	110000	15	BA000003_4	Continuation (5 of
c 846	103	6.6	348174	13	CR382399	Plasmidii	c 919	102	6.5	125828	12	AC135312	AC135312 Medicago
c 847	102.5	6.6	1527	4	AB048828	AB048828 Zeylanidi	c 920	102	6.5	128997	12	AC149473	AC149473 Medicago
c 848	102.5	6.6	1728	15	AY681274	Campyloba	c 921	102	6.5	145999	10	AF409138	AF409138 Goatpox v
c 849	102.5	6.6	2198	13	AF016033	Caenorhab	c 922	102	6.5	150509	10	AF409138	AF409138 Lumpy ski
c 850	102.5	6.6	3727	15	CLONTC	Di2739 C. butyricu	c 923	102	6.5	169794	12	AC004688	AC004688 Plasmidii
c 851	102.5	6.6	6289	2	AX599006	AX599006 Sequence	c 924	102	6.5	227066	12	AC006877	AC006877 Caenorhab
c 852	102.5	6.6	7291	15	PF211133	U40232	c 925	102	6.5	249943	13	AE014823	AE014823 Plasmidii
c 853	102.5	6.6	8066	13	PFU40232	U40232 Plasmidii	c 926	102	6.5	249995	13	AE014840	AE014840 Plasmidii
c 854	102.5	6.6	8913	10	AB198742	AB198742 Tonato sp	c 927	102	6.5	253924	13	AE014822	AE014822 Plasmidii
c 855	102.5	6.6	9289	2	AX767546	AX767546 Sequence	c 928	102	6.5	286680	13	AC024817	AC024817 Caenorhab
c 856	102.5	6.6	9295	15	AY573819	AY573819 Campyloba	c 929	102	6.5	298804	12	AC006911	AC006911 Caenorhab
c 857	102.5	6.6	10049	15	AF291051	AF291051 Carsonell	c 930	102	6.5	349980	2	AX344564	AX344564 Sequence
c 858	102.5	6.6	12214	15	AE010098	AE010098 Streptoco	c 931	101.5	6.5	922	13	PFAASN5514	M59473 Plasmidii
c 859	102.5	6.6	13616	15	AE006611	AE006611 Streptoco	c 932	101.5	6.5	1420	2	BD247497	BD247497 Promoters
c 860	102.5	6.6	16943	15	AE002153	AE002153 Streptoco	c 933	101.5	6.5	1420	2	AX026013	AX026013 Sequence
c 861	102.5	6.6	56153	2	AX251552	AX251552 Sequence	c 934	101.5	6.5	2004	13	AB189726	AB189726 Plasmidii
c 862	102.5	6.6	72325	13	AC025722	AC025722 Caenorhab	c 935	101.5	6.5	2425	4	AF315938	AF315938 Nepenthes
c 863	102.5	6.6	77835	4	PFMAL13P2_3	Continuation (4 of	c 936	101.5	6.5	2430	4	AF315884	AF315884 Nepenthes
c 864	102.5	6.6	92758	15	AE017332_8	Continuation (9 of	c 937	101.5	6.5	3913	15	S73676	S73676 neurotoxin
c 865	102.5	6.6	97405	15	AE017243_8	Continuation (9 of	c 938	101.5	6.5	5016	2	AX489296	AX489296 Sequence
c 866	102.5	6.6	106192	12	AC170582	AC170582 Medicago	c 939	101.5	6.5	5018	13	AC006797	AC006797 Caenorhab
c 867	102.5	6.6	108484	4	AC004238	AC004238 Arabidops	c 940	101.5	6.5	11075	12	AC156588	AC156588 Glomus in
c 868	102.5	6.6	110000	12	PFMAL13_21	Continuation (22 o	c 941	101.5	6.5	13440	2	AX191749	AX191749 Sequence
c 869	102.5	6.6	110000	10	PFMAL7P1_00	AL844506 Plasmidii	c 942	101.5	6.5	14009	13	PFCOMP1RB	X95276 Plasmidii
c 870	102.5	6.6	110000	12	PFMAL8P1_06	Continuation (7 of	c 943	101.5	6.5	36145	12	AL513224	AL513224 Drosophil
c 871	102.5	6.6	110000	15	AE001437_07	Continuation (8 of	c 944	101.5	6.5	92985	5	AL157378	AL157378 Human DNA
c 872	102.5	6.6	110000	15	CP000056_15	Continuation (16 o	c 945	101.5	6.5	97724	15	BA000021_6	Continuation (7 of
c 873	102.5	6.6	110000	15	CP000084_05	Continuation (6 of	c 946	101.5	6.5	110000	4	CR380957_03	Continuation (4 of
c 874	102.5	6.6	110000	15	CP000099_27	Continuation (28 o	c 947	101.5	6.5	110000	4	CR380957_04	Continuation (5 of
c 875	102.5	6.6	110000	15	CP000099_28	Continuation (29 o	c 948	101.5	6.5	110000	4	CR382135_11	Continuation (12 o
c 876	102.5	6.6	110000	15	CP000153_10	Continuation (11 o	c 949	101.5	6.5	110000	12	TANNA_02	Continuation (3 of
c 877	102.5	6.6	110000	15	AE009951_10	Continuation (11 o	c 950	101.5	6.5	110000	13	AC116984_0	AC116984 Dictyoste
c 878	102.5	6.6	110000	15	AE009951_11	Continuation (12 o	c 951	101.5	6.5	110000	13	CP000225_5	Continuation (6 of
c 879	102.5	6.6	110000	15	AE010299_52	Continuation (53 o	c 952	101.5	6.5	110000	13	CP000225_5	Continuation (6 of

c 953	101.5	6.5	110000	15	AE000783_01	Continuation (2 of
c 954	101.5	6.5	110000	15	AE001439_04	Continuation (5 of
c 955	101.5	6.5	110000	15	CP000061_0	CP000061 Aster yel
c 956	101.5	6.5	110000	15	CP000061_2	Continuation (3 of
c 957	101.5	6.5	110000	15	CP000123_04	Continuation (5 of
c 958	101.5	6.5	110000	15	AE000951_11	Continuation (12 of
c 959	101.5	6.5	110000	15	AE010300_06	Continuation (7 of
c 960	101.5	6.5	110000	15	AE010300_07	Continuation (8 of
c 961	101.5	6.5	110000	15	AE017243_4	Continuation (5 of
c 962	101.5	6.5	110000	15	AE017244_04	Continuation (5 of
c 963	101.5	6.5	110000	15	AE017332_4	Continuation (5 of
c 964	101.5	6.5	110000	15	AP008934_16	Continuation (17 of
c 965	101.5	6.5	110000	15	BA000021_1	Continuation (2 of
c 966	101.5	6.5	122890	4	AP005672	AP005672 Physcomit
c 967	101.5	6.5	127642	4	AC127020	AC127020 Medicago
c 968	101.5	6.5	133267	13	CP0000198	CP0000198 Drosophila
c 969	101.5	6.5	138465	4	AC148154	AC148154 Medicago
c 970	101.5	6.5	149475	12	AC175882	AC175882 Strongylo
c 971	101.5	6.5	162384	5	AC034236	AC034236 Homo sapi
c 972	101.5	6.5	177707	5	AC006265	AC006265 Homo sapi
c 973	101.5	6.5	192187	13	AC117072	AC117072 Dictyoste
c 974	101.5	6.5	216603	13	AC164195	AC164195 Bos tauru
c 975	101.5	6.5	250022	13	AE014824	AE014824 Plasmodiu
c 976	101.5	6.5	254436	13	AE014827	AE014827 Plasmodiu
c 977	101.5	6.5	299820	12	AC006803	AC006803 Caenorhab
c 978	101.5	6.5	299820	12	AC006871	AC006871 Caenorhab
c 979	101.5	6.5	315079	15	MPULM03	MPULM03 Mycoplasma
c 980	101.5	6.5	346939	15	EX842643	EX842643 Mycoplasma
c 981	101	6.5	1295	13	AF315277	AF315277 Plasmodiu
c 982	101	6.5	1524	4	AF344559	AF344559 Lophanthu
c 983	101	6.5	1779	2	AR678651	AR678651 Sequence
c 984	101	6.5	1839	2	AR678650	AR678650 Sequence
c 985	101	6.5	3199	15	HIU68399	HIU68399 Haemophilus
c 986	101	6.5	3318	2	AR707075	AR707075 Sequence
c 987	101	6.5	3318	2	AX392717	AX392717 Sequence
c 988	101	6.5	3572	10	RBL293984	RBL293984 Rice blac
c 989	101	6.5	3944	4	SCVGR134W	SCVGR134W S.cerevisia
c 990	101	6.5	4361	13	AF449199	AF449199 Plasmodiu
c 991	101	6.5	5663	13	CTU60196	CTU60196 Chlamydia m
c 992	101	6.5	9000	13	TETRRTRNA	TETRRTRNA Tetrahymena
c 993	101	6.5	9775	13	DQ011584	DQ011584 Plasmodiu
c 994	101	6.5	11357	15	AE002272	AE002272 Chlamydia
c 995	101	6.5	22897	4	AF008237	AF008237 Chlamydom
c 996	101	6.5	38062	13	CEFS6G4	CEFS6G4 Caenorhabdi
c 997	101	6.5	38454	15	AP008935	AP008935 Staphyloc
c 998	101	6.5	43380	13	CEY26D4A	CEY26D4A Caenorhab
c 999	101	6.5	47653	13	AY700145	AY700145 Polysphon
c1000	101	6.5	69663	13	AC024753	AC024753 Caenorhab

RESULT 1	AX934431	AX934431	AX934431	876 bp	DNA	linear	PAT 05-JAN-2004
LOCUS	AX934431	Sequence 8 from Patent WO02074942.					
DEFINITION	AX934431						
ACCESSION	AX934431						
VERSION	AX934431.1	GI:40641678					
KEYWORDS							
SOURCE	Campylobacter jejuni						
ORGANISM	Campylobacter jejuni						
REFERENCE	1						
AUTHORS	Gilbert, M. and Wakarchuk, W. W.						
TITLE	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics						
JOURNAL	Patent: WO 02074942-A 8 26-SEP-2002;						
FEATURES	National Research Council of Canada (CA)						
1..876							
1..876							
/organism="Campylobacter jejuni"							
/mol_type="unassigned DNA"							

/db_xref="taxon:197"							
1..876							
/note="unnamed protein product: Campylobacter							
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.							
jejunii O:19"							
/codon_start=1							
/transl_table=11							
/protein_id="CAF04169.1"							
/db_xref="GI:40641678"							
/translation="MKKVIILAGNGPSLKEIDYSRLPNDPFDVPRCNQFYFEDKYYLGGK							
CKAVFTPNFFEQYITLKLHIOQEYELIMCSNQAHLNENFVKTFDYDFPDA							
HLGYDFKQLKEFNAYFKFHEIYFNQRTTSGVCAVAIAGLYKEIYISGIDYQNGS							
SYADFTQENLLKLADPDKRSHYIGHSKNTDKALEFLKTYIKLYKILCPNLSLA							
NFIELAPNLNSNFIIOEKNNVTKDILLPSSEAYGKFSKNINPKIKIKENVYKLIKID							
LLRLPSDKHYFKGK"							

ORIGIN							
US-10-734-719-9 (1-291) x AX934431 (1-876)							
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20							
Db 1 ATGAAAAAGTATTATTGCTGGAAATGACCAAGTATTTAAAAAGAAATGATTTATTCAGG 60							
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40							
Db 61 CTACCAATGATTTGATGATTTAGATGATTAATCAATTTATTTTGAAGATAAATACTAT 120							
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60							
Db 121 CTGTGAAAAAATGCANAGCAGTGTTTACACCCCTAATTTCTTCTTGAGCAATACTAC 180							
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80							
Db 181 ACTTTAAACATTTAAATCAAAATCAAGATATGAGACCGAATAATATATGTGTCTAAT 240							
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100							
Db 241 TACACCAAGCTCATCTAGAAAATGAAATTTTGAAAAACATTTTACGATTTATTTTCT 300							
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120							
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360							
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140							
Db 361 TTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420							
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160							
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGTCA 480							
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180							
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540							
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200							
Db 541 GATGCTCGCACTATATCGGACATAGTAAAAATACAGATATATAAGCTTTAGAAATTTCTA 600							
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220							
Db 601 GAAAAAATCTTACAAATATAAACTATATGCTTATGCTTCTTAAATAGTCTTTTACCAATTTT 660							
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240							



Db	661	ATAGAACTAGCGCCAAATTTAAATTTCAATTTATCATACAGAAAAAATAACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATACATACCTTCTAGTGAGGCTTAAGGAAATTTTCAAAAAATATTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu	280
Db	781	AAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTTGATAAAGAATCTATTAAAGTTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGATATAAGCATTATTTCAGAGAAAA	873
RESULT 2			
LOCUS	BD249794	876 bp	DNA linear PAT 17-JUL-2003
DEFINITION	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics.		
ACCESSION	BD249794		
VERSION	BD249794.1	GI:33059564	
KEYWORDS	JP 2002535992-A/5.		
SOURCE	Campylobacter jejuni		
ORGANISM	Campylobacter jejuni		
REFERENCE	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.		
AUTHORS	1. Gilbert, M. and Wakarchuk, W.W.		
TITLE	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics		
JOURNAL	Patent: JP 2002535992-A 5 29-OCT-2002;		
COMMENT	NATIONAL RESEARCH COUNCIL OF CANADA		
	OS	Campylobacter jejuni	
	PN	JP 2002535992-A/5	
	PD	29-OCT-2002	
	PF	01-FEB-2000 JP 2000597438	
	PR	01-FEB-1999 US 60/118213 31-JAN-2000 US 06/495406 PI	
	MICHEL	GILBERT, WARREN W WAKARCHUK	
	PC	C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC	
	Campylobacter glycosyltransferases for biosynthesis of gangliosides and		
	ganglioside mimics		
	FH	Key	Location/Qualifiers
	FT	source	1. .876 /organism='Campylobacter jejuni'.
FEATURES	source		
	1. .876		
	/organism="Campylobacter jejuni"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:197"		
ORIGIN			
Alignment Scores:			
Pred. No.:	6.36e-190	Length:	876
Score:	1563.00	Matches:	291
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0
US-10-734-719-9 (1-291) x BD249794 (1-876)			
Qy	1	MetLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGATTATTATTCGTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	61	CTACCAATGATTTTGATGTATTTAGATGTATCAATTTTATTTTGAAGATAAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	121	CTTGTAATAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGAGCAATACTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAACATTTAATCCAAATCAAGAAATATGAGCGAAGCTAATATATGTGTCTTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATATTATTCCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTTAAA	360
Qy	121	PheHisGluIleTyrPheAsnGlnArGilleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCACGAAATTTTATTCATCAAGAAATTACCTCAGGGGTCTATATGTGTGAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAAATGATTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCCTATGCTTTTGATACCAACCAAGAAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCCTCGCACTATATCGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAAGAACTTACAAAAATAAACTATATTCCTTATGCTTAATAGTCTTTTAGCAAAATTTT	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAAATAACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATACATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu	280
Db	781	AAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAGAATCTATTAAAGTTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGATATAAGCATTATTTCAGAGAAAA	873
RESULT 3			
LOCUS	AR271703	876 bp	DNA linear PAT 10-APR-2003
DEFINITION	Sequence 8 from patent US 6503744.		
ACCESSION	AR271703		
VERSION	AR271703.1	GI:29703248	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 876)		
AUTHORS	Gilbert, M. and Wakarchuk, W.W.		
TITLE	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics		
JOURNAL	Patent: US 6503744-A 8 07-JAN-2003;		
	National Research Council of Canada; Ottawa;		
	CAX;		
FEATURES	Location/Qualifiers		
source	1. .876		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			



Pred. No.: 6.36e-190 Length: 876  
Score: 1563.00 Matches: 291  
Percent Similarity: 100.0% Conservatives: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR271703 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Dy 1 ATCAAAAGTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Dy 61 CTACCAATGATTGTTGATGTTTATGATGTTTATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Dy 121 CTGGTAAATAATGCAAGCAGTGTTTACACCCCTAATTTCTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Dy 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Dy 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Dy 301 GATGCTCATTTGGATATGATTTTAAACAATTAAGAAATTTAATGCTTATTATAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Dy 361 TTTTACCAAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Dy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Dy 481 TCTTAAGCTTTTGATATCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Dy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATAAAAAAGCTTTAGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Dy 601 GAAAAAATTTTCAAAATAAAAATCTATTTGCTTATGCTCTAATAGTCTTTTACGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Dy 661 ATAGACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAATTAATTAATTTT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Dy 721 AAAGATATATCATCATCTCTTAGTGAGCTTATGAGAAATTTTCAAAATTAATTAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Dy 781 AAAAAAATAAAAATTAAGAAATGTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Dy 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873

RESULT 4  
AR481785 876 bp DNA linear PAT 14-MAY-2004  
LOCUS AR481785  
DEFINITION Sequence 8 from patent US 6699705.

ACCESSION AR481785  
VERSION AR481785.1 GI:47243420  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert,M. and Wakarchuk,W.W.  
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides  
JOURNAL Patent: US 6699705-A 8 02-MAR-2004;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES Location/Qualifiers  
source 1..876  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores: 6.36e-190 Length: 876  
Pred. No.: 1563.00 Matches: 291  
Score: 100.0% Conservatives: 0  
Percent Similarity: 100.0% Mismatches: 0  
Best Local Similarity: 100.0% Indels: 0  
Query Match: 100.0% Gaps: 0  
DB: 2  
US-10-734-719-9 (1-291) x AR481785 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Dy 1 ATCAAAAGTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Dy 61 CTACCAATGATTGTTGATGTTTATGATGTTTATCAATTTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Dy 121 CTGGTAAATAATGCAAGCAGTGTTTTACACCCCTAATTTTATTTTGAAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Dy 181 ACTTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAACTAATTTATGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Dy 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Dy 301 GATGCTCATTTGGATATGATTTTAAACAATTAAGAAATTTAATGCTTATTATAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Dy 361 TTTTACCAAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Dy 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Dy 481 TCTTAAGCTTTTGATATCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Dy 541 GATCGCTCGCACTATATCGGACATAGTAAAAATACAGATAAAAAAGCTTTAGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Dy 601 GAAAAAATTTTCAAAATAAAAATCTATTTGCTTATGCTCTAATAGTCTTTTACGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Dy 661 ATAGACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAATTAATTAATTTT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Dy 721 AAAGATATATCATCATCTCTTAGTGAGCTTATGAGAAATTTTCAAAATTAATTAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Dy 781 AAAAAAATAAAAATTAAGAAATGTTTATTTACAAAGTTGATAAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Dy 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873

```
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPhelIleGlnGluLysAsnAsnTyrThr 240
|||
Db 661 ATAGAACTAGCCCAAAATTAAATTCAAATTTATCATACAGAAAAATAACTACACT 720
|||
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
|||
Db 721 AAGAGATATACATCATCTCTAGTGAGGCTTATGGAATAATTTCAAAAAATTAATTTT 780
|||
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysAspIleLeuArgLeu 280
|||
Db 781 AAAAAAATAAAAAATTAAGAAAAATCTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
|||
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
|||
Db 841 CCTAGTGATTAAGCAATTAATTTCAAGGAAAA 873
|||
RESULT 5
AR527384
LOCUS AR527384 876 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 8 from patent US 6723545.
ACCESSION AR527384
VERSION AR527384.1 GI:53914405
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1. (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Polypeptides having .beta.-1,4-GalNac transferase activity
JOURNAL Patent: US 6723545-A 8 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
Location/Qualifiers
source 1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 6.36e-190 Length: 876
Pred. No.: 1563.00 Matches: 291
Score: 100.0% Conserved: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 2
US-10-734-719-9 (1-291) x AR527384 (1-876)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
|||
Db 1 ATGAAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60
|||
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
|||
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTGGAAGATAAATCTAT 120
|||
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
|||
Db 121 CTGTGTAATAAATGCAAGCAGTGTTTACACCCCTAATTTCTTTGAGCAATACTAC 180
|||
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
|||
Db 181 ACTTTAAACATTTTAATCAAAATCAAGAATATGAGCCGAACTAATTTATGTGTTCTAAT 240
|||
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
|||
Db 241 TACACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTTATTTCT 300
|||
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
|||
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGATTTAATGCTTTATTTAA 360
|||
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
|||
```

```
Db 361 TTTACGAAATTAATTTCAATCAAGAATTAACCTCAGGGTCTATATGTGTGAGTAGGCC 420
|||
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
|||
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAAATTTGATTTTATCAAAATGGGTCA 480
|||
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
|||
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540
|||
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
|||
Db 541 GATCGCTCGCACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
|||
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220
|||
Db 601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATATGCTCTTTTAGCAAAATTTT 660
|||
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
|||
Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACTACT 720
|||
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
|||
Db 721 AAGATATATCTCATCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 780
|||
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
|||
Db 781 AAAAAAATAAAAAATTAAGAAAAATGTTTATTACAGTTGATAAAGATCTATTAAAGATTA 840
|||
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
|||
Db 841 CCTAGTGATTAAGCAATTAATTTCAAGGAAAA 873
|||
```

```
RESULT 6
AR609663
LOCUS AR609663 876 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 8 from patent US 6825019.
ACCESSION AR609663
VERSION AR609663.1 GI:56664963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1. (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W.W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 8 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
Location/Qualifiers
source 1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 6.36e-190 Length: 876
Pred. No.: 1563.00 Matches: 291
Score: 100.0% Conserved: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 2
US-10-734-719-9 (1-291) x AR609663 (1-876)
QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
|||
Db 1 ATGAAAAAAGTTATTATTGCTGGAATGACCAAGTTTAAAAAGAAATTTGATTTTCAAGG 60
|||
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
|||
```

```
Db 61 CTACCAATGATTTTGATGATTTAGATGUAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAATAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGATATGAGACCAAGCTAATATGTTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Qy 101 AsnAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAACCAACTTAAAGAAATTTAAATGCTTTATTTAAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnAlaValPheTyrThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACAGAAATTTATTTTCAATCAAGAAATTTACCTCAGGGTCTATATGTCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTAGCCCTGATTTTAAATAAT 540
Qy 181 AsnArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGACATATTCGGACATAGTAAATACAGATATTAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATAAAATATAATTTGCTTATGTCCTAATAGTCTTTTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATATCATACAAAGAAAAATAACTACACT 720
Qy 241 LysAsnIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTATACATCTTCTAGGAGCTTATGGAATTTTCAAAATAATTTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTAA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873
```

```
RESULT 7
AR689941
LOCUS AR689941 876 bp DNA linear PAT 12-SEP-2005
DEFINITION Sequence 8 from patent US 6905867.
ACCESSION AR689941
VERSION AR689941.1 GI:74471949
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl
transferase activity
JOURNAL Patent: US 6905867-A 8 14-JUN-2005;
National Research Council of Canada, Ottawa;
CAX; Location/Qualifiers
FEATURES
source 1..876
```

```
ORIGIN
Alignment Scores: 6.36e-190 Length: 876
Pred. No.: 1563.00 Matches: 291
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 2
```

US-10-734-719-9 (1-291) x AR689941 (1-876)

```
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTTGCTGGAAATGGCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCAAGCTAATATGTTCTTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAAATTTTACGATTTATTTTCT 300
Qy 101 AsnAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACCAACTTAAAGAAATTTTAAATGCTTTATTTAAA 360
Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTACCAAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGATACAAAGAAATTTATCTTTGGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAAAATCTAGCCCTGATTTTAAAAAT 540
Qy 181 AsnArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCGACATATTCGGACATAGTAAATACAGATATTAAGCTTTTGAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAATAAAATATAATTTGCTTATGTCCTAATAGTCTTTTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTTCAATTTTATATCATACAAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTATACATCTTCTAGGAGCTTATGGAATTTTCAAAATAATTTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAAAGATTAA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATTTATTTCAAGGAAAA 873
```







```

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 5351 TCTTATGCTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 5410
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200
Db 5411 GATCGCTCGACTATATCGGCATAGTAAATAACAGATATAAAGCTTTAGATTCTTA 5470
Qy 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 5471 GAAAAAATACAAAATAAAATATATTGCTTATGCTCTATAGTCTTTTACGAAATTTT 5530
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 5531 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAAGAAAAATACTACACT 5590
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 5591 AAAGATATATCATACCTTCTAGTGAGCTTATCGAAAAATTTTCAAAAAATATTAAATTT 5650
Qy 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeuArgLeu 280
Db 5651 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAGATCTATTAAAGATTA 5710
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 5711 CCTAGTATATAAGCATTATTTCAAGGAAAA 5743

RESULT 11
LOCUS AY422197 24425 bp DNA linear BCT 02-FEB-2004
DEFINITION Campylobacter jejuni putative ABC-type transport protein, putative
UDP-glucose-4-epimerase, hypothetical protein, putative
heptosyltransferase I, putative lipid A biosynthesis
acyltransferase, putative two-domain glycosyltransferase, putative
glycosyltransferase, putative
beta-1,4-N-acetylgalactosaminyltransferase, putative
alpha-1,3-galactosyltransferase, putative
alpha-2,3-sialyltransferase, putative sialic acid synthase,
putative UDP-N-acetylgalactosamine 2-epimerase/N-acetylmannosamine
synthase, putative CMP-Neu5Ac synthetase, putative
acetyltransferase, putative glycosyltransferase, putative
heptosyltransferase II, putative phosphonopentose isomerase 1,
putative ADP-heptose synthase, putative
ADP-glyceromannohexose-6-epimerase, putative phosphatase, putative
periplasmic cytochrome C, small hypothetical hydrophobic protein,
and putative cation-transporting ATPase genes, complete cds; and
putative transcription termination factor gene, partial cds.
AY422197
ACCESSION AY422197.1 GI:40217896
VERSION
KEYWORDS
SOURCE
ORGANISM Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 24425)
AUTHORS Li, J., Wakarchuk, W.W. and Endtz, H.P.
TITLE Evidence for Acquisition of the Lipooligosaccharide Biosynthesis
Locus in Campylobacter jejuni GB11, a Strain Isolated from a
Patient with Guillain-Barre Syndrome, by Horizontal Exchange
Infect. Immun. 72 (2), 1162-1165 (2004)
14742567
JOURNAL 2 (bases 1 to 24425)
PUBMED Gilbert, M., Karwaski, M.-F., Cunningham, A. and Brochu, D.
AUTHORS Direct Submission
TITLE Submitted (24-SEP-2003) Institute for Biological Sciences, National
JOURNAL Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
OR6, Canada
FEATURES Location/Qualifiers
source 1..24425
/organism="Campylobacter jejuni"

```

```

/mol_type="genomic DNA"
/strain="GB11"
/serotype="HS.2"
/db_xref="taxon:197"
complement(21..1715)
/codon_start=1
/transl_table=11
/product="putative ABC-type transport protein"
/protein_id="AAR82866.1"
/db_xref="GI:40217897"
/translation="MLKKLFFILSKEDKNFLFFLLVFSVFISFIETFAISLWMPFTIL
ASDFYDRNKYLISLKEYLNIPVEIIVYFGVLIVFYVFRALLNAVYFHLARFSK
GRYHAIAVYKSKFLINYEKPTQNSQSEILKSTITGVNLSSTFSLMSSEIFV
LLVALMLINKTIFLFSIFWLVNAFLIKLSPILKAGVRRSEAMKNFTEILNTN
LNNEFKIKLTKEDGVLSLFAQSEAFKANTNESVAAPRIYLEGICFCLVFTV
FLVKNESDISGILSTISIFVLALYRLMPSANRIITSYHDLLYHSSLDIIYONLROE
EENGEKLSNQELKICNLSFGYEGKILFKNLNLNKKGBKIAFIGSGCGKSTLV
DLIIGLLPKEGQILIDQELNANNNTKYNKIGYIPQNIYLFNDSIAKNIITFGDAD
BEKLNRVIKQANLEHFIKNLPQGVQTKVGGGNSLGGQKQIAIARALYLPEDMLV
DEATSAIDQSEAKIMDBIYKISKDKTWIIIAHRLSTITQCDKVYRLEHGLKEEK"
complement(1709..2695)
/codon_start=1
/transl_table=11
/product="putative UDP-glucose-4-epimerase"
/protein_id="AAR82867.1"
/db_xref="GI:40217898"
/translation="WKILISGGAGVIGSHLTROFLKTDHEICVLNLSKSGKIAIEDL
QKIRAFKFFEQDLSDFQVKALFEREKFDALVHPAASEVFSFMSQNPVKLYNNNTVNT
SMNPTEFCKILRYFNAGACMDYTLGQRYPKATLLIKVAAECAAGKDKLFIQDDY
VDFKVELAPRRAGDPSVLISDASLKIRNLTSQPKYDLDLELIKSAFDEWEKQC"
complement(2753..3583)
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAR82868.1"
/db_xref="GI:40217899"
/translation="WILAKLPKRLKMAKNBYICVFDCESVPDVELIKTLGFEQSD
LEVSUKALQWKEQSEFLPLFYHKIISICAVLSDNFGKFIKVNKIDQNKEMLIEN
FFYFTEYEPKLVSFNGKNDPMLVLKALYNLKAATYLDTSFMLNTDNFNAKLAKQDVFS
HCDLIESIGSGRGIKLDTLSMVGKGVHGVDEVMKLFYENLEKIIYECESVL
NTYMLFLKVELIKANVSEEDYVDFLSYMRDLFEKKKSDRSYTEVFAKACESISIKVRS
"
3613..4641
/codon_start=1
/transl_table=11
/product="putative heptosyltransferase I"
/protein_id="AAR82869.1"
/db_xref="GI:40217900"
/translation="WKIALVRLSALGDIIQSAVVLQFIKNFKDIEIHWVDEKFEGI
LKNHPLIDKLYALPLKDKKILKSLKARKNNNAVLDLQGLIKSAIVSLRSNN
FGDKNSLSEFAHNFYKNQLELDNENFVRYLSLTSFMLNTDNFNAKLAKQDVFS
VDENLKQNLNKLKIDKNEKILIHVGSSVENKIYPTKTLAILCKLLINFPQTKIWL
AWGNVKEFEFAKELVNLGSDIETHTELPKFKLELMFTKMDLIIGNDSGPTHLAF
ALNKASITIFGATPSYRNAFQTHINKIIDTGKLIKQNAKIDKSDFCITRIEEDIFKL
AKGLUNEK"
4631..5518
/codon_start=1
/transl_table=11
/product="putative lipid A biosynthesis acyltransferase"
/protein_id="AAR82870.1"
/db_xref="GI:40217901"
/translation="MKNSDRIYLSLYILKFPVFTMPDCILHFLALIVARIAFLNKK
HRKIDNALAKRPILFTTHYGNWEILSLAVAAGYGVISVGGKLSKSEVWYILSOSR
ENFLIDALQICFPQYTKERDKLSKIYENFAQFGIDCLQNTTKKILANKVFIN
TQPIELIDKKGIRQMSALKKEPAGLITLDDCCVENESVRLKPNKENVYQMGASL
IAQRNRALIPYATKEGKFCIEFFKAKDSQNASLEELTLQAQSCSEMIKKRPNEY
FFFFRRFASVNEEIIYKGAK"
5515..6768
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"

```

```
/protein id="AAR82871.1"
/db_xref="GI:40217902"
/translation="MNLKQISVIIIVTKNAEOTLLCLNSLKDFDEIILINNESSONTL
KIANEPKDFANLYIYHNAFTGFGALKNLALSYAKNDMLISIDADADEVLENCEIKELKN
LKQEDNIIALSRKNLYKGWMIKAGWMPDYVLIRPKNKFRFNDNLVHESVLPSNA
KKIYIKNGLSHYSDISHLDRQWYSSMAKQNHKKSGVLKANLRAFVFFRNYF
LNGFLYGYKGFIIISVCSALGTFYFKMLYELQROKPKTCALIIITYNQERLKLVLDS
KNLAFIYVLIADDSGDEPTARLIEEYQKDFPCPKLHIWOEDRGFKLSRNNKTIK
NADSEYIIVIDGDMLEKDFIKHELEFAQRKLFLOGSRVILNKKESEBELNKDDVRII
FNKDFPKSKNSFLAKIYFISLKKR"
7064..8233
/codon_start=1
/transl_table=1
/product="putative glycosyltransferase"
/protein id="AAR82872.1"
/db_xref="GI:40217903"
/translation="MKKIGVPIIYNVKEYIURELDSVINQTYNNLEIILVNDGSTDE
HSLNAKEYTLKDRITLFDKNGGLSSARNIGIEYFSGEYKLNKTHIENSIEIF
QLDGNPNYIYKAVKSOAFNNEKDLTFTYPSIDYIIFLDSNDYKLNCEIECVIRM
KNVDLWFDHDTVEDNTKNKHKTRMEIFDFKKECIITPKKEYANRALSVGSRDIFSFG
WNGMIDNFLKQIKLKFNFINEDIHFGIILFASANKIYVLSQKLYLCRLRANSISN
HKKITPCANSEYFKDIYETFGENAKEARNYLDKAASRVITALKLIEFFKQDKQENALA
IKETFLPCYAKKALMIKKFKDPLNLKESQLVLKPFIIQTLPYIDIWKFQWQIKNI"
8303..9346
/codon_start=1
/transl_table=1
/product="putative
beta-1,4-N-acetylglucosaminyltransferase"
/protein id="AAR82873.1"
/db_xref="GI:40217904"
/translation="WLFOSYFVKIICLPIPFKRIRHKIKTFLNLIQORDKIDSYLPK
KTLVQINKYNNEDILKNKAIIREGHKGFFNYDEKSKDPKSPINPMAIRVQNEAITL
KASLELPAIORGVINDCTDSEIILEFCQTPSFIPIKYPIEYIQIQNKSEEN
KLYSYNNYVAFIPEKWEIIKIDVDHIYDAKKLYKSYFIPKNKDVVSYSRVDIHYFN
DNFLCNDNNIILEPGDCLLNNYNLKWKEVLDIRNNMKWKATQKQSFSGNIHSLE
OLKYKHLRILFHTLENNYHFFPLKXRAQDIYKYNWISIEEFKKFYLQINNHKIEPSMI
SKETLKKIILFLF"
complement(9332..10237)
/codon_start=1
/transl_table=1
/product="putative beta-1,3-galactosyltransferase"
/protein id="AAR82874.1"
/db_xref="GI:40217905"
/translation="WFKSIIILPTYNVQYITARATESCINQTFKDIIEIIVDDCGNDN
SINTAKESKDKRIKIIHNENKUGLLRAREGVKANSPIYIMFLDPDIDYELNACEE
CIKILDEQEDVLDVFNFAIVESNVIYSYKFPNSGYFKSEFKKIIIAKNLYWTWNG
KLIRKLYLEAPSLRLKQKVKINNAEDVLLYPMLSQAQKIAYMNCNLYHYVPPNNS
ICNTNKEVLNKNNTQELQVLNLRQNYILNKYCSVLVYLKLYLIQIYIKIRTKLM
VTLAKINILTKILPKYKFLKQC"
10326..11201
/codon_start=1
/transl_table=1
/product="putative alpha-2,3-sialyltransferase"
/protein id="AAR82875.1"
/db_xref="GI:40217906"
/translation="MKKVIIAAGNGPSLKEIDYSRLPNDPVPFCNQVFEDKYYLGKK
CKAVFYTNFFEQYITLKHQIYETELTMCNSYNOAHLENENFYKTFYDYPDA
HLGDFPFKQELFNAPFHFHYFNQITSGVMCAVAIALGYEIKIYLSGIDFYQNGS
SVAFDTKQENLKLAPDPKDRSHVGHSGKNTDIKALEFLEKTYKIKIYCLCPNSLLA
NFIELAPLANSFNIIQENKNTYTKDILIPSSRAYGFKSNINFPKKIKIKENYVYKLIKD
LLRLPSDIKHYFKGK"
11202..12242
/codon_start=1
/transl_table=1
/product="putative sialic acid synthase"
/protein id="AAR82876.1"
/db_xref="GI:40217907"
/translation="WKEIKIQNIISSEKAPLVVPEIIGINHNGSLLELAKIMVDAFSA
GAKTIKHOTHIVEDMSKAARKVIPGNAKISYEIMOKCALDYKDELAKYTEKLG
VYLGTFPRAGANLEDMGVAFKIFSGECNNYPLIKHIAAFKPMIVSTGMNSTESI
KPTVKILLNDIIPVFLMTNLVPTPHNLVRLNMLELKKKFSQWVGLDTHDNLAC
LGAVALGCVLHERHTDSMDRSGPDIVCSMDTKALKELIMQSEQWAINRGNESKAA
KQEQVTIDFAFASVVSIKDIKKGEVLSMDNIVWRRPGLGGISAVEFENILGKKALRDI
```

```
CDS
ENDTQLSYEDFA"
12239..13357
/codon_start=1
/transl_table=11
/product="putative UDP-N-acetylglucosamine
2-epimerase/N-acetylmannosamine synthase"
/protein id="AAR82877.1"
/db_xref="GI:40217908"
/translation="MKKILFITGTRADYSIKSLMYRVQNSSEPELYIFATGMHLNKN
FGYTVKELVKNGFKNIVFYNIDKYQTDKALATTIDGFSRYVNLKPDLLIVHGDRI
EPFLAAIVGALNNILVAHIEGSEISCTIDDSLRAHISKLAHILHIANDEFAKRLMOLG
EDEKSIPIIGSPDLELNDNKNLSLSEAKTYIDINENYALLMFHPVTEITSIKNQAD

Alignment Scores:
Pred. No.: 1 58e-188 Length: 24425
Score: 1563.00 Matches: 291
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 15 Gaps: 0

US-10-734-719-9 (1-291) x AY422197 (1-24425)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
|||||
Db 10326 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGG 10385

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
|||||
Db 10386 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTGGAAGATAAATACAT 10445

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
|||||
Db 10446 CTTGGTAAAAAATGCAAGCAGTGTTTTACACCCCTAATTTCTTTTGGACAATACATAC 10505

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
|||||
Db 10506 ACTTTAAACATTTAATCCAAATCAAGATAATGAGACCAAGCAATATTATGTGTTCTAAT 10565

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValLysThrPheTyrAspTyrPhePro 100
|||||
Db 10566 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTCCT 10625

QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
|||||
Db 10626 GATGCTCATTTGGATATGATTTTAAACAACCTTAAAGAAATTAATGCTTATTTTTAAA 10685

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
|||||
Db 10686 TTTACGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 10745

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
|||||
Db 10746 ATAGCCTAGGATACAAAGAAATTTATCTTCGGAAATGATTTTATCAAAATGGGTCA 10805

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
|||||
Db 10806 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 10865

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
|||||
Db 10866 GATCGCTCGACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGATTCTTA 10925

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
|||||
Db 10926 GAAAAAATCTACAAATAAAATATATTGCTTATGCTCTAATAGCTTTTAAAGCTTTT 10985

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
|||||
Db 10986 ATAGAACTAGCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAAATAACTACT 11045

QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
|||||
```



```

Db      11046 AAGATATACCTACCTCTTAGTGAGGCTATTGCAAAATTTTCAAAAAATTTTAAATTTT 11105
Qy      261  LysIleLysLysLysGluAsnValTyrTyrLysLeuLysLysAspLeuLeuArgLeu 280
        |||||
Db      11106 AAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 11165
Qy      281  ProSerAspIleLysHisTyrPheLysGlyLys 191
        |||||
Db      11166 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 11198

RESULT 12
LOCUS   AF167344                24437 bp    DNA        linear    BCT 23-JAN-2004
DEFINITION Campylobacter jejuni HS:19 lipooligosaccharide biosynthesis locus,
partial sequence.
ACCESSION AF167344
VERSION   AF167344.2
KEYWORDS  GI:41058381
SOURCE    Campylobacter jejuni
ORGANISM  Campylobacter jejuni
          Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
          Campylobacteraceae; Campylobacter.
REFERENCE
AUTHORS  Gilbert,M., Brisson,J.R., Karwaski,M.F., Michniewicz,J.,
          Cunningham,A.M., Wu,Y., Young,N.M. and Wakarchuk,W.W.
TITLE     Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
          Identification of the glycosyltransferase genes, enzymatic
          synthesis of model compounds, and characterization of nanomole
          amounts by 600-mhz (1)h and (13)c NMR analysis
          J. Biol. Chem. 275 (6), 3896-3906 (2000)
JOURNAL  10660542
PUBMED   2 (bases 1 to 24437)
REFERENCE Gilbert,M., Cunningham,A.-M., Karwaski,M.-F., Michniewicz,J.,
          Wu,Y., Young,N.M. and Wakarchuk,W.W.
AUTHORS  Direct Submission
TITLE     Submitted (08-JUL-1999) Institute for Biological Sciences, National
          Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
          0R6, Canada
JOURNAL  3 (bases 1 to 24437)
PUBMED   Gilbert,M.
AUTHORS  Direct Submission
TITLE     Submitted (22-JAN-2004) Institute for Biological Sciences, National
          Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
          0R6, Canada
REMARK   Nucleotide and amino acid sequences updated by submitter
COMMENT  On Jan 22, 2004 this sequence version replaced gi:6969981.
FEATURES
         Location/Qualifiers
             1..24437
                 /organism="Campylobacter jejuni"
                 /mol_type="genomic DNA"
                 /strain="ATCC 43446"
                 /serotype="HS:19"
                 /db_xref="ATCC:43446"
                 /db_xref="taxon:197"
                 complement(21..1715)
                 /codon_start=1
                 /transl_table=1
                 /product="putative ABC-type transport protein"
                 /protein_id="AAR99161.1"
                 /db_xref="GI:41058384"
                 /translation="MLKKLFILSKEDKNFLFLLVPSVFSFTEAISLVMPFITL
                 ASDFSPDRNKYLSLEYLNIPVEIIVFGVGLIVFYVFRALLNAYIFHLLARFSK
                 GRHYATAYKFSFLNTNIEKTKNQSEILKSTIGEVNLSNIMSSFLMLMSIFVY
                 LLLVLMLLNKTITFLSLPMVNAFLIKLSPILIKKAGLRREEMKNPFELTNTN
                 LNNFKFTKLATKDGVLSPKAQSEAFSKANITNESVAAVPRIYLEGIGFCLVFIIV
                 FVLKNESDLSGITLSTISIPVLALYRLMPSANRIITSYHLLVYHSSLDIIYONLROE
                 EENLEGKJFNQELKCNLSFGYEGKGHLPKNLNNIKKEKIAFTGEGCGKSTLV
                 DLTIGLPKPEGQLDKQELNASNANKYQKIGIPQNIYLFNDSIAKNITFGDAVD
                 EEKLSVKIKQANLEHFIKNLPQGVQTKVGGGNSLGGQKQIAIALALVLEPILVL
                 DEATSLDITQSEAKIMDEIVKISDKTMIIIAHLRLTITQCDKVRLEHGLKKEK"
                 complement(1709..2695)
                 /codon_start=1
                 /transl_table=1

```

```

/product="putative UDP-glucose-4-epimerase"
/protein_id="AAR99162.1"
/db_xref="GI:41058385"
/translation="MKLLISGAGYIGSHTLRQFLKTDHEICVLDNLKSGSKIAIEDL
KILRAFKEFDLSDFOQKALFEREKDAIVHFASIEVFESMQNPLKIYINNTVNT
QNLLETCTGTGNKFIIFSTAAATYGEPTQVPSVETSPAPINPYSKLMSEVLRWD
SWANPEFKCILRYFNVAGACMDYLGORYPRATLLIKVAACCAAGKDKLFIFGDDY
DTKDGTCTIRDFIHVDDISSAHLAALDYIQENESNVFVGYGHGFSVKEVIEAMKRVSG
VDFVELAPRAGDPSVLISDASKIRNLTSWQPKYDDLELICKSAFDEWKQC"
complement(2753..3583)
/inference="non-experimental evidence, no additional
details recorded"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAR99163.1"
/db_xref="GI:41058386"
/translation="MILAKFKRRLKMAKNEGYICVDFCESVPDVELIRKTLFGESD
LEVLKALQWQEQSGSEFLPLPYHKIISICAVLSDNFGFIKVNKIDGQNEKEMIEI
FNFTENYEPKLVSPNGNFDMPVLVLRALKYNLKAATYLDTSQDKNNYKTFSELK
HCDLLESLSNGRGIKLDTLCSMVLPGKYDVHGVEMKLFYENELEKLYEYESDVL
NTYMLFLKYLKIANVSEEDYVDFLSYMRDFLREKKSRSYSYEVFAKASEISEKRS
"
3613..4641
/note="formerly ORF-01; similar to Helicobacter pylori
RfaC"
/codon_start=1
/transl_table=1
/product="heptosyltransferase I"
/protein_id="AAF34138.2"
/db_xref="GI:41058382"
/translation="MKIALVRLSALGDIIQSAVLQFIKFKDKDIEIHWFDVKFEGI
LKNHPLDKLYALPLKDKKILKSLKILKARKNNYNAVLDLQGLIKSAVSRILSRNN
FGDKNSLKSEFAHNFYQKLELDYNNVFRYLSLTSFMNTDFNAKLAFKQDVFS
VDENLQULNNKLKLDKNEKNILIHVGSVENKIYPTKLAILCKLINFQOTKIWL
AWGNKYEYFAKEVLNLSGIDETHIELAPFNLEELMAFTKMDMLIIGNDSGTHLAF
ALNKASITIFGATPSYRNAFQTHINKIIDTGKKIQNAKHIDKSDFCITRIEEDIFKL
AKGLLNK"
4631..5518
/note="formerly ORF-02; similar to Helicobacter pylori
WaaM"
/codon_start=1
/transl_table=1
/product="acyltransferase"
/protein_id="AAF34139.1"
/db_xref="GI:6969984"
/translation="MKNSDRIYLSLYILKFFVTFMPDCILHFLALIVARIAFHLNKK
HRKIINTNLQICFPQYTKOKERDKLSLKIYENFAQFQIDCLQNTQTKELINKVNFN
ENFLIDALAKRPIITTAHYGNWEILSLAYAAKYCAISIVGKKLSEVMYVYLSQSR
TQFDIELIDKKGIGIOMLSALKERALGILTQDCVNESESRLKFFNKENVYOMGASL
IAQRSNALIIPYATKEGKGCIEFFKAKDSQNASLEELTLVQAQSCSEMIKKRPWEY
FFFHRRFASPYNEEIIYKGA"
5515..6771
/note="formerly ORF-03; similar to Neisseria meningitidis
LgtP"
/codon_start=1
/transl_table=1
/product="glycosyltransferase"
/protein_id="AAF34140.1"
/db_xref="GI:6969985"
/translation="MNLKQISVIIIVKNAEOTLLECLNSIKQDFEILLNNSSDNTL
KIAREFKDPAFLYIYHNAFGGALKNALSYAKNDWILSDADVENECIKELKN
LKLQEDNIIALSRLKNLYGEMTKACGWPDYVLRINFRNFRPNLNLVHESLVFNPA
KKIILKNGLKHYSYKIDSHLIDMQYYSLLMAKQNIHKSGVLKANRLKRWFTFRNYF
LKNGFLGYKGFIIISVCALGTFPKYKLYELQKQPKTCALIIITYNQKRLKLVLD
SKVNLARLPNVLIAADGSKEDTARLIBEYQKDFPCPLKHIWOEDGFKLSKSRNKT
KNASEYIIVIDGMILEKDFIKHELFQAQRKLFLQSGRVILNKKSEELNKDDYRI
IFNKADPKSSKNSNPLAKIFYLSKKR"
7067..8236
/note="formerly ORF-04; similar to Streptococcus
pneumoniae Cps14J"
/codon_start=1
/transl_table=1

```



Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

## REFERENCE

## AUTHORS

Gilbert,M. and Wakarchuk,W.W.

## TITLE

Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics

## JOURNAL

Patent: WO 02074942-A 13 26-SEP-2002;

## FEATURES

National Research Council of Canada (CA)

Location/Qualifiers

1..873

/organism="Campylobacter jejuni"

/mol\_type="unassigned DNA"

/db\_xref="taxon:197"

1..->873

/note="unnamed protein product; Campylobacter

alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.

jejuni O:36"

/codon start=1

/transl table=11

/protein\_id="CAFO4171.1"

/db\_xref="GI:40641683"

/translation="MKKVIAGNPSLKEIDYSLPNDVDFRCNQFYFEDKYYLGLKK

CKTVFFPNFFQYTLKHLIQNEYETELIMCSNYNOAHLENENFVKTFDYFPDA

HLGVDFPKQENLNAFKFHEIYNQRIITSGVMCAVAIALGYKEIYLSGLDFYQNGS

SYAFDFKQENLLKLPDKNDRSHVIGHSKNTDILKALEFLEKTKIKLYCLCPNSLLA

NFIELAPLNSNFIQEKNNYTKDILIPSEAYGKFSKNINFKKIKIKENVYYKLKD

LLRLPSDIKHFKGK"

## ORIGIN

Alignment Scores:

Pred. No.: 2,06e-189 Length: 873

Score: 1559.00 Matches: 290

Percent Similarity: 99.7% Conservative: 0

Best Local Similarity: 99.7% Mismatches: 1

Query Match: 99.7% Indels: 0

DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AX934436 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20

Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40

Db 61 CTACCAAAATGATTGATGATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

Db 121 CTGTGTAATAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80

Db 181 ACITTTAAACATTTAATCCAAATCAAGAAATATGAGACCGAATTAATTTATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100

Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTGTAAAAACTTTTACGATTATTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

Db 301 GATGCTCATTTGGGATATGATTTTTTAAACCACTTAAAGAAATTTAATGCTTATTATAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

Db 361 TTTACCAAAATTTAATTCATCAAGAAATTAACCTCAGGGGTCTATATGTCGATAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrIleuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180

Db 481 TCTTATGCTTTTGTATACCAAAACAGAAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

Db 541 GATGCTCCTACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220

Db 601 GAAAAAACTTACAAAAATAAACTATATTGCTTATGCTTATGCTTATGCTTATGCTTATGCT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnLeuTyrThr 240

Db 661 ATAGAACTAGCCCAAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAAATTAACACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

Db 721 AAGATATACCTACCTCTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTATAATTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280

Db 781 AAAAAAATAAAATTAAGAAAAATGTTTATTACAAAGTTGATTAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291

Db 841 CCTAGTATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 14

AR481787

LOCUS

DEFINITION

SEQUENCE 13 from patent US 6699705.

ACCESSION

AR481787

VERSION

AR481787.1

GI:47243422

KEYWORDS

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 873)

Gilbert,M. and Wakarchuk,W.W.

Campylobacter glycosyltransferases for biosynthesis of gangliosides

and ganglioside mimics

Patent: US 6699705-A 13 02-MAR-2004;

National Research Council of Canada; Ottawa;

CAN;

FEATURES

Location/Qualifiers

1..873

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 2,06e-189 Length: 873

Score: 1559.00 Matches: 290

Percent Similarity: 99.7% Conservative: 0

Best Local Similarity: 99.7% Mismatches: 1

Query Match: 99.7% Indels: 0

DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR481787 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20

Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40

Db 61 CTACCAAAATGATTGATGATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

Db 121 CTGTGTAATAAATGCAAAACAGTGTGTTTACACCCCTAATTTCTTTGAGCAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80

Db 361 TTTACCAAAATTTAATTCATCAAGAAATTAACCTCAGGGGTCTATATGTCGATAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrIleuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180

Db	181	ACTTTAAACATTTAAATCCAAATCAAGAATATGAGACCGAACTAAATTTATGGTGTCTAAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTATTTCTCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAAATGCTTATTTTAA	360
Qy	121	PheHisGluLeuTyrPheAsnGlnArgileThrSerGlyValTyrMetCysAlaValala	140
Db	361	TTTTCACGAAATTTATTCATCAAGAATTACCTCAGGGTCTATATGTGTGCAGTAGCC	420
Qy	141	IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTGATTTTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProCaspPheLysAsn	180
Db	481	TCTTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATCGCTCACACTATATCGGCACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLeLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAIAAAACTTTACAAAAATAAAACTATATTGCTTATGTCCTAATAGTCTTTTAGCAAAATTTT	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlyLeuAsnAsnTyrThr	240
Db	661	ATAGAACTAGCGCCAAATTTAAATTTCAAATTTTATCATACAAGAAAAATAACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAGATATACTCATACCTTCTAGTAGGCGTTATGGAAAAATTTTCAAAAAATATTAAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAATAAANAATTAAGAAAAATGTTTTATTACAAGTTGATAAAGATCTATTAGATTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTGAATAAAGCACTATTTTCAAAGGAAAA	873

RESULT	15
AR527386	
LOCUS	AR527386
DEFINITION	Sequence 13 from patent US 6723545.
ACCESSION	AR527386
VERSION	AR527386.1 GI:53914407
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 873)
TITLE	Gilbert,M. and Wakarchuk,W.W.
JOURNAL	Polypeptides having .beta.-l,4-GalNAc transferase activity Patent: US 6723545-A 13 20-APR-2004; National Research Council of Canada; Ottawa;

		/mol_type="Genomic DNA"	
ORIGIN			
Alignment Scores:			
Pred. No.:	873		
Score:	2.06e-189		
Percent Similarity:	1559.00		
Conservative:	290		
Best Local Similarity:	99.7 %		
Mismatches:	1		
RESULTS		AR609665      873 bp      DNA	
LOCUS		Sequence 13 from patent US 6825019.	
DEFINITION		ACCESSION	
AR609665		AR609665	
VERSION		GI:56664965	
KEYWORDS		SOURCE	
UNKNOWN.			

ORGANISM Unknown.  
Unclassified.  
1 (bases 1 to 873)  
REFERENCE Gilbert,M. and Wakarchuk,W.W.  
AUTHORS Polypeptides having .beta.-1,3-galactosyl transferase activity  
TITLE Patent: US 6825019-A 13 30-NOV-2004;  
JOURNAL National Research Council of Canada; Ottawa;  
CAX;  
FEATURES Location/Qualifiers  
source 1..873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,06e-189 Length: 873  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR609665 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATCTAT 120  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60  
Db 121 CTTCGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTAAATCCAAAAATCAAGAATATGAGACCGCACTAAATATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCT 300  
Qy 101 AspalahisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACGAAATTTTATTTTCAATCAAGAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATACCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTCA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATCTTCAAAAATAAAATATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAAATAACTTACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

Db 721 AAAGATATACTATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATATTATTAATTT 780  
Qy 261 LysLysLysLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280  
Db 781 AAAAAATATAAAATTTAAAGAAATGTTTATTACAGTTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 17  
AR689943  
LOCUS AR689943  
DEFINITION Sequence 13 from patent US 6905867.  
ACCESSION AR689943  
VERSION AR689943.1 GI:74471951  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 873)  
AUTHORS Gilbert,M. and Wakarchuk,W.W.  
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl  
JOURNAL transferase activity  
Patent: US 6905867-A 13 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES Location/Qualifiers  
source 1..873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,06e-189 Length: 873  
Score: 1559.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.7% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR689943 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40  
Db 61 CTACCAATGATTTTATGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATCTAT 120  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60  
Db 121 CTTCGTAATAAATGCAAAACAGTGTTTTACACCCCTAAATTTCTTTTGAGCAATCTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTAAATCCAAAAATCAAGAATATGAGACCGCACTAAATATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTTCT 300  
Qy 101 AspalahisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTTATTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTACGAAATTTTATTTTCAATCAAGAATTTACCTCAGGGGCTATATGTCGAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATACCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 240  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 200  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTCA 220  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 240  
Db 601 GAAAAAATCTTCAAAAATAAAATATATGCTTATGCTCTAATAGTCTTTTACGAAATTTT 260  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 280  
Db 661 ATAGACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAAATAACTTACACT 300  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 320

```

Db      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480
Qy      161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db      481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT 540
Qy      181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541 GATGGCTCACACTATATCGGCATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy      201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLysLeuAlaAsnPhe 220
Db      601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy      221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleLysGlnGluLysAsnThrTyrThr 240
Db      661 ATAGAAGCTAGCGCCAAATTTAAATTTTATCATACAGAAAAAATAAATACACTACACT 720
Qy      241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721 AAAGATATACCTACATACCTCTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTT 780
Qy      261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781 AAAAAATAAAATAAAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAGATTA 840
Qy      281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 18
AR691839
LOCUS      AR691839      873 bp      DNA      linear      PAT 13-SEP-2005
DEFINITION Sequence 13 from patent US 6911337.
ACCESSION AR691839
VERSION   AR691839.1 GI:74479550
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 873)
AUTHORS  Gilbert,M. and Wakarchuk,W.W.
TITLE    Nucleic acids encoding .beta.-1,4-GalNAc transferase
JOURNAL  Patent: US 6911337-A 13 28-JUN-2005;
        National Research Council of Canada; Ottawa;
        CAX;
FEATURES  Location/Qualifiers
           source      1..873
                    /organism="unknown"
                    /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      2,06e-189      Length:      873
Score:          1559.00      Matches:      290
Percent Similarity: 99.7%      Conservative: 0
Best Local Similarity: 99.7%      Mismatches: 1
Query Match:    99.7%      Indels:      0
DB:             2      Gaps:      0

US-10-734-719-9 (1-291) x AR691839 (1-873)
Qy      1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1 ATGAAAAAGTTATTATTGCTGGAAATGACCAAGTTTAAAGAAATTTGATTATTCAGG 60
Qy      21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61 CTACCAATGATTGTTGATGTTTATGATGTAATCAATTTATTTTGAAGATAAATACTAT 120
Qy      41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

```

```

Db      121 CTTGTGTAATAAATGCAAAACACAGTGTTTTACACCCCTAAATTTCTTCTTTGAGCAATACTAC 180
Qy      61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181 ACTTTAAACATTTTATCCAAATCAAGATATAGACCGAAGTAATATTATGTTCTTAAT 240
Qy      81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241 TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAAACTTTTACGATTATTTTCTCT 300
Qy      101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301 GATGCTCATTTGGGATATGATTTTAAACAACTTTAAAGAAATTTAATGCTTATTTTAAA 360
Qy      121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361 TTTACAGAAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTTGCGAGTAGCC 420
Qy      141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTGATTTTATCAAAATGGGTCA 480
Qy      161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db      481 TCTTATGCTTTTGATACCAAAACAAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy      181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTTAGAAATTTCTA 600
Qy      201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      601 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 660
Qy      221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db      661 ATAGAAGCTAGCGCCAAATTTTAAATTTTATCATACAGAAAAAATAACTACACT 720
Qy      241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721 AAAGATATACCTACATACCTCTCTAGTGAGGCTTATGGAAATTTTCAAATAATATTAAATTTT 780
Qy      261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781 AAAAAATAAAATAAAAGAAATGTTTATTACAAAGTTGATAAAAGATCTATTAGATTA 840
Qy      281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 19
AF401528
LOCUS      AF401528      12390 bp      DNA      linear      BCT 03-JAN-2002
DEFINITION Campylobacter jejuni O:36 heptosyltransferase I (waaC) gene,
           partial cds; lipid A biosynthesis acyltransferase (htrB), putative
           two-domain glycosyltransferase, putative glycosyltransferase,
           truncated beta-1,4-N-acetylglucosaminyltransferase (cgtA-I),
           beta-1,3-galactosyltransferase (cgtB), alpha-2,3-sialyltransferase
           (cst-II), sialic acid synthase (neuB1), putative
           N-acetylglucosamine-6-phosphate 2-epimerase (neuCl),
           beta-1,4-N-acetylglucosaminyltransferase (cgtA-II), CMP-Neu5Ac
           synthetase (neuAl), putative acetyltransferase, and putative
           glycosyltransferase (waaV) genes, complete cds; and
           heptosyltransferase II (waaF) gene, partial cds.
ACCESSION AF401528
VERSION   AF401528.1 GI:15718482
KEYWORDS  Campylobacter jejuni
SOURCE    Campylobacter jejuni
ORGANISM  Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
           Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 12390)
AUTHORS  Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E.,

```

TITLE Michniewicz,J., Cunningham,A.M. and Wakarchuk,W.W.  
The Genetic Bases for the Variation in the Lipo-oligosaccharide of  
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF  
STALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE  
J. Biol. Chem. 277 (1), 327-337 (2002)  
JOURNAL 11689567  
PUBMED  
REFERENCE 2 (bases 1 to 12390)  
AUTHORS Gilbert,M., Michniewicz,J., Karwaski,M.-F., Cunningham,A. and  
Wakarchuk,W.W.  
TITLE Direct Submission  
JOURNAL Submitted (22-JUL-2001) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada

FEATURES Location/Qualifiers  
source 1..12390  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="ATCC 43456"  
/serotype="O:36"  
/db\_xref="ATCC:43456"  
/db\_xref="taxon:197"  
<1..348  
/gene="waaC"  
CDS <1..348  
/gene="waaC"  
/codon\_start=1  
/transl\_table=11  
/product="heptosyltransferase I"  
/protein\_id="AAL05984.1"  
/db\_xref="GI:15718483"  
/translation="FAKEVNLGSDTHIELAPKFNLELMAPTQMDLLIGNDSGP  
THLAPLNKASITTFGTPSYRNAFQTHINKIITGKKIQNAKHIDKSDCITRIBE  
DIFPLAKGLNEK"  
338..1225  
/gene="htxB"  
338..1225  
/gene="htxB"  
/codon\_start=1  
/transl\_table=11  
/product="lipid A biosynthesis acyltransferase"  
/protein\_id="AAL05985.1"  
/db\_xref="GI:15718484"  
/translation="WKSNDRIYLSLYILKFFVTFMPDCILHPLALIVARIAPHLNKK  
HRKINTNLQICFPQYQKEDKLSKIYENFAQFGIDCLONQTTKEILKNYFN  
ENFLIDALAKRPIIFTAHYGNWELLSLAYAAGALSIVGKLGSEVMYELISQSR  
TQFDIELIDKGGIRQMLSKALKERALGILTDQCVENESVRLKFNKENVYQMGASL  
IAQSNALIIPVAYKEGGKFCIEFFRAKDSQNASBELTLYQAQSCSEEMIKRPWEY  
FFEHRRFASVNEEIVKGAK"  
1222..2769  
/note="orf3"  
/codon\_start=1  
/transl\_table=11  
/product="putative two-domain glycosyltransferase"  
/protein\_id="AAL05986.1"  
/db\_xref="GI:15718485"  
/translation="WNLKQISVILIIIVNABQTLLECLNSLKDFEILLNNESIDNLT  
NIANEFKDFANLYIYNHAFGFGALKNLALSIAKNDWILSIDADEVLENCIKELKN  
LELEQNDIILSRNLKLGKWIWACGWSPDYVLRIFNKNFTFRNDNLVLSLVPNV  
KKHLKNGLKHVAFRDISHLINKMQYSSYLWAKQNIHKKSGVLKANLRAFWTFRRNVF  
LKNGLFYQCGQTSVNCALGTFFKMKLYELOKQPKTCALIIITYNOKERKLVLVD  
SVKNLTFPNVLVADGSKEDTARLIEEYQKDFPCSLKHIWQDEGFRAAKRNNAI  
RVSKSEYIILIDGMULEKQDFKNHFBAKRVILQGSRVILNKNESEILKNDQDYSL  
AFSKDFKSYNSFLAKIYIKSIDQKQFNTKEIIRKSGKTCNMSFFKADFDLQDGN  
ENFTGWGREDSEFVARFLFNKGIIFRLKFKAIAYHIVHEENSKOMLESNHQIYLDITIK  
NKKTSWR"  
2760..3944  
/note="orf4"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAL05987.1"  
/db\_xref="GI:15718486"  
/translation="WEIKHEENRCSYTNIIQRRRIIFRMFNSVINQTYTNLILVNDG

gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS

STDNSLNIAREYTLKDERFILFDKKNGLSSARNVGLIEYFSGEYKLNKNTQHIKENS  
LIFQNDVNPYNIYKAYKSSQAFNNEKDLNFTYPNIDYIIFLSDDYWEINCEIEEC  
VIRMKNDVLMFHDHCTYEDNIKNKHKTRMEIIFDFKKECIITPEYANRALSVGSRD  
ISGWNIGIDFNLKIKLKFNLINELIHFGILLFASANKYIILSKLVLCRURAN  
SISNHDKKITKANVSEYFKDIYETFGENAKKAYLKAASRVTTALKLIEFPKQDKE  
NALAIKETFLFCYAKKMIKKFKDPLNLEQQLVLTKPFIQTKLPIYDIWIKFWQKIKK  
IN"  
3988..4248  
/gene="cgtA-I"  
3988..4248  
/gene="cgtA-I"  
/note="orf5-I; premature translation stop due to a one bp  
deletion"  
/codon\_start=1  
/transl\_table=11  
/product="truncated  
beta-1,4-N-acetylglactosaminyltransferase"  
/protein\_id="AAL05988.1"  
/db\_xref="GI:15718487"  
/translation="MLFQSYFVKIICLIPERKIRHKIKTFLKNIQRDKIDSYLPK  
KILIQINKYNNEDLIKLKAITGGIKIDILIMKMLKIQNLL"  
complement (5016..5921)  
/gene="cgtB"  
complement (5016..5921)  
/gene="cgtB"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,3-galactosyltransferase"  
/protein\_id="AAL05989.1"  
/db\_xref="GI:15718488"  
/translation="MFKISIIILPTYNVQYIARAIESCINOTFKDIEIIVDDCGNDN  
SINTAKYSKKDKRIKIHNKKNLGLLRARYEGVKVANSPIVIMFLDPPDYLELNACEE  
CIKLLDQDENVLVEFNALVESNVIYSKGFDFNSGFSYKKEFVKI IAKNCLYWTMVG  
KLRRKLYLEAFSLRELEKDVKNMAEDVLLIYPMLSQAOKIAYNVCNLYHYVPNNNS  
ICNTKNEVLVKNNIQELQVLNLRQNYILNKYCSVLVLIKILLYIQIYKIRKTLKM  
VTLAKINILTKILFKYKFLKQC"  
6010..6885  
/gene="cst-II"  
6010..6885  
/gene="cst-II"  
/codon\_start=1  
/transl\_table=11  
/product="alpha-2,3-sialyltransferase"  
/protein\_id="AAL05990.1"  
/db\_xref="GI:15718489"  
/translation="MKKVVIIAGNPSLKEIDYSRLPNDPFRQCNQFYEDKYILGKK  
CTVFTYPNFFEQYTYTLKLIQOEYETELIMCSYNOAHLENFVKTFYDYPDA  
HLGYDFPKQKFEFAYKFHEIYFNQRTSGYMCVAIAIAGYKIEIYSLGDFYQNGS  
SYAFDTQENLLKLAPDFKNDRSHVIGHSKNTDIKALEFLKTYKIKLYCLCPNSLLA  
NFIELAPLNSNFIIQEKNNYTKDILIPSEAYGKFSKNINFPKIKIKENVYVKLIKD  
LLRLPSIKHYFKGK"  
6886..7926  
/gene="neuB1"  
6886..7926  
/gene="neuB1"  
/codon\_start=1  
/transl\_table=11  
/product="sialic acid synthase"  
/protein\_id="AAL05991.1"  
/db\_xref="GI:15718490"  
/translation="MKEIKIQNIIESEKAPLVPEIIGINHNGSLAKIMVDAAFSA  
GAKLIKQTHIVEDEMSKAKVIPGNAKISITYEMOKCALDYKDELALKEYTKLGL  
VYLTSPFRAGANRLDMGVSAFKIGSGECNNYPLIKHIAAFKPKMIVSTGWNISIESI  
KPTVYKILLDNEIPFVLMHTNLYPPTHNLVRNLNMLELKEPSCMWGUSDHTIDNLAC  
LGAVALGCVLERHFTDSMHRSGPDIVCSMDTKALKELIIQSEQWAIMKGNESKCA  
KOEQVTFDAFASVVSIIKIKKGEVLSDMNDIIVWVRPGLGGISAAEFENILGKALRDI  
ENDTQLSYDEFA"  
7923..9047  
/gene="neuC1"  
7923..9047  
/gene="neuC1"  
/codon\_start=1  
/transl\_table=11

```

/product="putative N-acetylglucosamine-6-phosphate
2-epimerase"
/protein_id="AAL05992.1"
/db_xref="GI:15718491"
/translation="MKKILFVTGTRADYSKIKSLMYRVONSSEFELYIPATGMHLSKN
FGYTYKEYKNGFNKNIFYFINYDKYQTDKALATTIDGFSRYVNLKPDLIIVHGDRI
EPLAAIIVGALNLIIVHIEGGEISGTTDDSKLHAI SKLAHILHVNDFAKRLMQLG
EDEKSIPIIGSPDLLELNDNKISLSEAKKYDINVENVALLMFHPVTEITSIKNQAD
NLVXALIQSNKYLIVIPNNDLGFELILQSVLEELKNRPFKLPSLRFEYRITLLKNA
DFIGNSSCILKEALYLITAGILVGSRONRGLNGENTILKVNANSDEILKAINTHKKQ
DLFTKLEILDSKSLFFELYSUGBFFKLSTQKVFQDIKHA"
9040. .9987
/gene="cgtA-II"
9040. .9987
/gene="cgtA-II"
/gene="cgtA-II"
/note="orf5-II"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylglucosaminyltransferase"
/protein_id="AAL05993.1"
/db_xref="GI:15718492"
/translation="WLKILISLYKRYISKLVLDNEHPIKENKNYKGGKHGFFDFD
EKADVKSPMPGFRVKNKALTLRVLSLEILPALQRIIAYNDCDSEELILEFC
KOYNFIAQENKALCYPRINFIILNGNIYVONSNGYVFIGGDDLLIKRRNSPTERR
YKSFYRIDQENKALCYPRINFIILNGNIYVONSNGYVFIGGDDLLIKRRNSPTERR
VSKSQWIDPKGLIEELYSEQOVLQGVKILQAPLLOWHFPALKRNDYQOYLILS
LEEFQAFHRKSKRAKKIDFAMLRKPVIEQILKKFOGEIK"
9984. .10649
/gene="neuA1"
9984. .10649
/gene="neuA1"
/codon_start=1
/transl_table=11

```

## Alignment Scores:

```

Pred. No.:      2,678-188      Length:      12390
Score:          1559.00      Matches:      290
Percent Similarity: 99.7%      Conservative: 0
Best Local Similarity: 99.7%      Mismatches: 1
Query Match:      99.7%      Indels:      0
DB:              15          Gaps:      0

```

US-10-734-719-9 (1-291) x AF401528 (1-12390)

```

Qy 1  MetlysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6010 ATGAAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG 6069
Qy 21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6070 CTACCAAAATGATTTTGTATGATTTATAGATGTAATCAATTTTATTGGAAGATAAATCTAT 6129
Qy 41  LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6130 CTTGTAAAAAATGCAAAACAGTGTTTACACCCCTAATTTCTTTTGGAGCAATACTAC 6189
Qy 61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6190 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAACTAATATATGTGTTCTAAT 6249
Qy 81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6250 TACACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCT 6309
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6310 GATGCTCAITTTGGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 6369
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6370 TTTCAGAAATTTATTTCAATCAAGAAATTACCTCAGGGGTCTATATGTGTCAGTAGCC 6429
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

```

```

Db 6430 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTATTTATCAAAATGGGTCA 6489
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6490 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 6549
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6550 GATCGCTCACATATATCGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 6609
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6610 GAAAAAACTTACAAATAAAACTATATTGCTTATGCTTAATAGTCTTTTAGCAAAATTTT 6669
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6670 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT 6729
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6730 AAAGATATATCTATACCTTCTAGTAGGCTTATGGAAATTTTCAAAAAATATTAATTTT 6789
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6790 AAAAAAATAAAATTAAGAAATGTTTATTACAAGTTGATAAAAGATCTATTAAAGATTA 6849
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6850 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 6882

```

RESULT 20  
AX934434  
LOCUS  
DEFINITION  
Sequence 11 from Patent WO02074942.  
ACCESSION  
AX934434  
VERSION  
AX934434.1 GI:40641680  
KEYWORDS  
Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
Campylobacteraceae; Campylobacter.

REFERENCE  
1  
AUTHORS  
Gilbert,M. and Wakarchuk,W.W.  
TITLE  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
JOURNAL  
Patent: WO 02074942-A 11 26-SEP-2002;  
National Research Council of Canada (CA)  
FEATURES  
Location/Qualifiers  
source  
1. .873  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1. ->873  
/note="unnamed protein product; Campylobacter  
alpha-2,3/alpha 2,8-sialyltransferase II (CstII) from C.  
jejuni O:4"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAF04170.1"  
/db\_xref="GI:40641681"  
/translation="MKKVIAGNGLPSLKEIDYSRLPNDFVFRCNQYFPEDKYVLGKK  
KCAVYTFPFFFEQYTYLKHQIYQYETELINCYNQAHLENENFKVTFYFPDA  
HIGVYDFKOLFNFAYFKFEIYFNQITSGVYMCVAIALGYKEIYLSGIDFYNGS  
SVAFDTKQNLKLAPDPKNDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLA  
NFIEIAPNLNSNFILOEKNNYTKDILLPSSEAYGKFSKNINFKKIKIKENVYVKLKD  
LURLFSIDIKHYFKGR"

## ORIGIN

```

Alignment Scores:      3,718-189      Length:      873
Pred. No.:            1557.00      Matches:      290
Percent Similarity:  99.7%      Conservative: 0

```



Best Local Similarity:	99.7%	Mismatches:	1
Query Match:	99.6%	Indels:	0
DB:	2	Gaps:	0
US-10-734-719-9 (1-291) x AX934434 (1-873)			
Qy	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	1	ATGAAAAAGTTATTATTCGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG	60
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr	40
Db	61	CTACCAATGATTTGATGATTTAGATGTAATCAATTTTATTTGAAGATAAATACTAT	120
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr	60
Db	121	CTTGCTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTCTTTTGAGCAATACCTAC	180
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	181	ACTTTAAACAATTTAAATCCAAATCAAGAAATATGAGCCGAACTAATTAIGTGTCTAAT	240
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	241	TACAACCAAGCTCATCTAGAAAATGAAATTTTGTAAAAACTTTTACGATTTATTCCT	300
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	301	GATGCTCATTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAAATGCTTTTAA	360
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	361	TTTCAAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTCAGTACC	420
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	421	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA	480
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
Db	481	TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAAACTAGCCCTGATTTTAAAAAT	540
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	541	GATGCTCACACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTCTA	600
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	601	GAATAAACTTACAAAATTAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAGATTA	660
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr	240
Db	661	ATAGACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTACACT	720
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
Db	721	AAAAGATATACATACCTCTAGTGGGCTTATGGAAATTTTCAAAAAATTTAAATTTT	780
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280
Db	781	AAAAAATTAATAAATAAAGAAATGTTTATTTACAGTTGATAAAGATCTATTAGATTA	840
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
Db	841	CCTAGTATATAAGCATTTATTTCAAGGAAAA	873
RESULT 21	AR481786	Sequence 11 from patent US 6699705.	
LOCUS	AR481786	873 bp	DNA linear PAT 14-MAY-2004
DEFINITION	AR481786		
ACCESSION	AR481786		
VERSION	AR481786.1	GI:47243421	
KEYWORDS			

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 873)
TITLE	Gilbert M. and Wakarchuk, W.W.
JOURNAL	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics
FEATURES	Patent: US 6699705-A 11 02-MAR-2004; National Research Council of Canada; Ottawa; CAX;
source	Location/Qualifiers 1..873 /organism="unknown" /mol_type="genomic DNA"
ORIGIN	
Alignment Scores:	
Pred. No.:	3.71e-189
Score:	1557.00
Percent Similarity:	99.7%
Best Local Similarity:	99.7%
Query Match:	99.6%
DB:	2
US-10-734-719-9 (1-291) x AR481786 (1-873)	
Qy	1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db	1 ATGAAAAAGTTATTATTGCTGAAATGGACCAAGTTTAAAGAAATGATTATTCAGG 60
Qy	21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db	61 CTACCAAAATGATTTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATACTAT 120
Qy	41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db	121 CTTGCTAAAAAATGCAAGCAGTGTTTTACACCCCTGGTTTCTTTCTTTGAGCAATACTAC 180
Qy	61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db	181 ACTTTAAACAATTTAATCCAAATCAAGAAATATGAGCCGAACTAATTAATGTGTCTTAAT 240
Qy	81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db	241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACTTTTACGATTAATTTTCCT 300
Qy	101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db	301 GATGCTCATTGGGATATGATTTTTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTTAA 360
Qy	121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db	361 TTTCAAGAAATTTATTTCATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420
Qy	141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db	421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTCATTTTTATCAAAATGGGTCA 480
Qy	161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db	481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTTAAACCTAGCCCTGATTTTAAAAAT 540
Qy	181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db	541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy	201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db	601 GAAAAAACTTACAAAATAAAACTATATTGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660
Qy	221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db	661 ATGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACACT 720



Db 121 CTTGGTAAAAAATGCAAGAGAGTGTTCCTTACACCCCTGGTTCCTTGTGACATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACITTTAAACATTTAATCCAAAATCAAGATATGAGACCGAATCAATATATGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGNAATTTTGTAAAACTTTTACGATTATTTTCCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTAAACAATTAAGAAATTAATGCTTATTTTAAA 360  
Qy 121 PheHisGluLysTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCGAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAACAGTCTTTTACGCAAAATTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAATTTTATCATCAAGAAATAAATACTACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATCTATCATCTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATATTAAATTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 24  
AR689942  
LOCUS AR689942 873 bp DNA linear PAT 12-SEP-2005  
DEFINITION Sequence 11 from patent US 6905867.  
ACCESSION AR689942  
VERSION AR689942.1 GI:74471950  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 873)  
AUTHORS Gilbert,M. and Wakarchuk,W.W.  
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl  
transferase activity  
JOURNAL Patent: US 6905867-A 11 14-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES  
source Location/Qualifiers  
1..873  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
RESULT 25  
AR691838

Alignment Scores:  
Pred. No.: 3 71e-189 Length: 873  
Score: 1557.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.6% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR689942 (1-873)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAAGTTATTTATTTGCTGGAAATGCACCAAGTTTAAAAAGAAATTTGATTATTCAAGG 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAAAATGATTTTGTAGATGTTATCAATTTTATTTTGAAGATAAATACTAT 120  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTTGTAAAAAATGCAAGACGATGTTTACACCCCTGGTTCCTTTGAGCAATACTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATTTATGTGTTCTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCCT 300  
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTITTAACAACCTTAAAGAAATTTAATGCTTATTTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTTCAAGAAATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAAAAATAGCCCTGATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAATAAACTATATTTGCTTATGCTTAAACAGTCTTTTACGCAAAATTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCCAAAATTTAAATTTCAATTTTATCATCAAGAAATAAATACTACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAGATATCTATCATCTCTAGTGAGCTTATGGAAAAATTTTCAAAAAATATTAAATTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAAATAAAATTTAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

LOCUS AR691838 873 bp DNA linear PAT 13-SEP-2005  
DEFINITION Sequence 11 from patent US 6911337.  
ACCESSION AR691838  
VERSION AR691838.1 GI:74479549  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 873)  
AUTHORS Gilbert, M. and Wakarchuk, W.W.  
TITLE Nucleic acids encoding .beta.-1,4-GalNAc transferase  
JOURNAL Patent: US 6911337-A 11 28-JUN-2005;  
National Research Council of Canada; Ottawa;  
CAX;  
FEATURES Location/Qualifiers  
source 1..873  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.71e-189 Length: 873  
Score: 1557.00 Matches: 290  
Percent Similarity: 99.7% Conservative: 0  
Best Local Similarity: 99.7% Mismatches: 1  
Query Match: 99.6% Indels: 0  
Gaps: 0  
DB: 2  
US-10-734-719-9 (1-291) x AR691838 (1-873)  
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAGATTATTATGCTCGAATGACCAAGTTTAAAGAAATGATTATTCACAGG 60  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAATGATTGATGATTATAGATGTAATCAATTTATTTTGAAGATAAATACAT 120  
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTTGTGTAATAAATGCAAGCAGTGTTTACACCCCTGGTTCTCTTTCGAGCAATAC 180  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCAAAATCAAGAATATGAGCCGAATTAATGTTCTAAT 240  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTTCT 300  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTAAA 360  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTTCAGAAATTTATTTCATCAAAAGAAATTTACCTCAGGGGCTATATGTGTCAGTAG 420  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGG 480  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTATGCTTTGATACCAACAAGAAATCTTTTAAACCTAGCCCTGATTTTAAAAAT 540  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 541 GATCGCTCACACTATATCGACATAGTAAAAATACAGATATAAAAAGCTTTAGAAATTT 600  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 601 GAAAAAATTTACAAAAATAAACTATATTGCTTATGCTTAAACAGTCTTTTAGCAAAATTT 660

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
DB 661 ATAGAAGTACGCGCAAAATTTAAATTTCAAAATTTATCATCAAGAAAAATACTACTACT 720  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 780  
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAAATAAAATTAAGAAATGTTTATTCAAGTTGATAAAGATCTATTAAAGATTA 840  
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873  
RESULT 26  
AF215659 11474 bp DNA linear BCT 02-JAN-2001  
LOCUS  
DEFINITION Campylobacter jejuni lipooligosaccharide biosynthesis locus,  
partial sequence.  
ACCESSION AF215659  
VERSION AF215659.1 GI:12004275  
KEYWORDS  
SOURCE  
ORGANISM Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE 1 (bases 1 to 11474)  
AUTHORS Gilbert, M., Michniewicz, J. and Wakarchuk, W.W.  
TITLE The LOS biosynthesis locus of Campylobacter jejuni O:4  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 11474)  
AUTHORS Gilbert, M., Michniewicz, J. and Wakarchuk, W.W.  
TITLE Direct Submission  
JOURNAL Submitted (12-DEC-1999) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A  
0R6, Canada  
FEATURES Location/Qualifiers  
source 1..11474  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="ATCC 43432"  
/serotype="O:4"  
/db\_xref="ATCC:43432"  
/db\_xref="taxon:197"  
<1..360  
/notes="ORF-01; similar to RfaC from Helicobacter pylori"  
/codon\_start=1  
/product="heptosyltransferase I"  
/protein\_id="AAG43973.1"  
/db\_xref="GI:12004276"  
/translation="KEYEFAKEVLNLSGIDETHIELAPKESLELMFTKMDLIIGN  
DSGPHLAFALNKASITIFGATPSYRNAFTHINKIIDAGKKIQNAKHIDKSDCFITR  
IEEDFDIFKLAKGLLNK"  
350..1237  
/notes="ORF-02; similar to WaaM from Helicobacter pylori"  
/codon\_start=1  
/product="acylttransferase"  
/protein\_id="AAG43974.1"  
/db\_xref="GI:12004277"  
/translation="MKNSDRIYLSLVYILKFFVTFMPECILHFLALIVARTFTFYLNNK  
HKIDINLIQICFPQYTKERDKLSLKIYENFAQFGIDCLQNONQTTKEKILMYKYNFIN  
ENLLDALALKRPIIFTHYHGNWEILSLYAAKYGAISIVGKKLSEIMYIILSONR  
TFQDELIDKKGIQMLKKGKRGILTDQDCVEKESVRLKPFKNVNYQMGASL  
IAQRNALIPVYAYEGGKFCIEFFKAKDSQNASLEELTYQAQSCSEEMIKRPWEY  
FFHRRFASYNBEIYKGA"  
1234..2490  
/notes="ORF-03; similar to LgtF from Neisseria  
meningitidis"



Alignment Scores:		4, 47e-188	Length:	11474
Pred. No.:		1557.00	Matches:	290
Score:		99.7%	Conservative:	0
Percent Similarity:		99.7%	Mismatches:	1
Best Local Similarity:		99.6%	Indels:	0
Query Match:		15	Gaps:	0
DB:				
US-10-734-719-9 (1-291) x AF215659 (1-11474)				
Qy	1	MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20	
Db	6048	ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG	6107	
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40	
Db	6108	CTACCAATGATTTTGGTGTATTTAGATGTATCAATTTTATTATTGAAGATAAATACTAT	6167	
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60	
Db	6168	CTTGTAATAAATGCAAGCAGTGTTTTACACCCCTGGTTCTCTTTGAGCAATACTAC	6227	
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80	
Db	6228	ACTTTAAACATTTAATCCAAAATCAAGATATGAGCCGAATTAATATGTGTCTAAT	6287	
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100	
Db	6288	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAACTTTTACGATTATTTCCT	6347	
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120	
Db	6348	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	6407	
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140	
Db	6408	TTTCAGAAATTTATTCAATCAAGAATTAACCTCAGGGTCTATATGTGTGCAGTAGCC	6467	
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160	
Db	6468	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA	6527	
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180	
Db	6528	TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTGATTTTAAAAAT	6587	
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaIleuGluPheLeu	200	
Db	6588	GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTTGAATTTCTA	6647	
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220	
Db	6648	GAAAAAATTACAAATAAAACTATATTGCTTATGCTTAACAGCTTTTAGCAATTTT	6707	
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnThrThr	240	
Db	6708	ATAGAACCTAGCGCCCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAATAACTACACT	6767	
Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260	
Db	6768	AAAGATATACTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATTTAATTTT	6827	
Qy	261	LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu	280	
Db	6828	AAAAAATAAAATTAAGAAAAATTTTATACAGTTGTATAAAGATCTATTAGATTATTA	6887	
Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291	
Db	6888	CCTAGTGATAAAGCATTTATTTCAAGGAAAA	6920	
RESULT 27 AX934427				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
CDS				
AX934427				
Sequence 4 from Patent WO02074942.				
AX934427				
AX934427.1				
GI:40641674				
Campylobacter jejuni				
Campylobacter jejuni				
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;				
Campylobacteraceae; Campylobacter.				
1				
Gilbert,M. and Wakarchuk,W.W.				
Campylobacter glycosyltransferases for biosynthesis of gangliosides				
and ganglioside mimics				
Patent: WO 02074942-A 4 26-SEP-2002;				
National Research Council of Canada (CA)				
Location/Qualifiers				
1..876				
/organism="Campylobacter jejuni"				
/mol_type="unassigned DNA"				
/db_xref="taxon:197"				
1..876				
/note="unassigned protein product; bifunctional				
alpha-2,3/alpha 2,8-sialyltransferase Campylobacter				
sialyltransferase II (cstII) from C. jejuni serotype O:10				
(ORF 7a of lipooligosaccharide (LOS) biosynthesis locus)"				
/codon_start=1				
/transl_table=11				
/protein_id="CAP04167.1"				
/db_xref="GI:40641675"				
/translation="MKKVIAGNPSLKEIDYRLPNDFVRCNQFYFEDKYYLGGK				
FKAVFPGLPFEQVYTLKHLIQOBYETELIMCSNVNQAHLENENFVKFYDFPDA				
HGVDFPKLKEFNAYFKFHELYLNDRIITSGVVMCAVAIALGKVEIYLSGIDFYQNGS				
SVADTKQNLKLAPDFKNDRSHYIGHSKNTDIKALEFLEKTKYIKLYCLCPNSLLA				
NFLEAPLNSNFIIQEKNNYTDILIPSEAYGKFSKNINFKIKIKENIYYKLKD				
LRLPSDIKHYPKGK"				
ORIGIN				
Alignment Scores:				
Pred. No.:	1, 91e-185	Length:	876	
Score:	1528.00	Matches:	285	
Percent Similarity:	98.3%	Conservative:	1	
Best Local Similarity:	97.9%	Mismatches:	5	
Query Match:	97.8%	Indels:	0	
DB:	2	Gaps:	0	
US-10-734-719-9 (1-291) x AX934427 (1-876)				
Qy	1	MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20	
Db	1	ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATGATTATTCAAGG	60	
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40	
Db	61	CTACCAATGATTTTGGTGTATTTAGATGTATCAATTTTATTATTGAAGATAAATACTAT	120	
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60	
Db	121	CTTGGTAAAAAATTCAAAGCAGTATTTTTCAATTCCTGGTCTCTTTTGTGAACCAATCTAC	180	
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80	
Db	181	ACTTTAAACATTTAATCCAAAATCAAGAAATATGAGACCAACTAATATATGTGTCTAAT	240	
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100	
Db	241	TACAACCAAGCTCATCTAGAAATGAAATTTTGTAAAACTTTTACGATTATTTCCT	300	
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120	
Db	301	GATGCTCATTTGGGATATGATTTTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA	360	
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140	

```

Db      361  TTTCCAGAAATTTATCTCAATCAAAGAAATTAACCTCAGGAGTCTATATGTGTGCAGTAGCT 420
Qy      141  ILAlaIeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTCATTTTATCAAAATGGGTCA 480
Qy      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db      481  TCTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTGGCTCTCGATTTTAAAAAT 540
Qy      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTTACAAAATAAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660
Qy      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db      661  ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAAGAAAAATTAACACTACT 720
Qy      241  LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721  AAAGATATATCTATACCTCTTAGTGAGCTTATCGAAATTTTCAAAAAATATTAATTTT 780
Qy      261  LysLysIleLysLysLeuGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781  AAAAAATAAAATTAAGAAAAATATTTATTAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841  CCTAGTATATAAGCAATATTTTCAAAGGAAAA 873

RESULT 28
BD249792 876 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Campylobacter glycosyltransferases for biosynthesis of
gangliosides and ganglioside mimics.
ACCESSION BD249792
VERSION BD249792.1 GI:33059562
KEYWORDS JP 2002535992-A/3.
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert M. and Wakarchuk W. W.
TITLE Campylobacter glycosyltransferases for biosynthesis of
gangliosides and ganglioside mimics
JOURNAL Patent: JP 2002535992-A 3 29-OCT-2002;
NATIONAL RESEARCH COUNCIL OF CANADA
COMMENT OS Campylobacter jejuni
PN JP 2002535992-A/3
PD 29-OCT-2002
PF 01-FEB-2000 JP 2000597438
PR 01-FEB-1999 US 60/118213,31-JAN-2000 US 06/495406 PI
MICHIL GILBERT WARREN W WAKARCHUK
PC C12N15/09,C12N9/10,C12N9/88,C12N15/00 CC
Campylobacter glycosyltransferases for biosynthesis of CC
gangliosides and
CC ganglioside mimics
FT Key Location/Qualifiers
FT source 1..876
FT 1..876 /organism='Campylobacter jejuni'.
FEATURES
source Location/Qualifiers
1..876
/organism='Campylobacter jejuni'
/mol_type='genomic DNA'
/db_xref='taxon:197'

ORIGIN
Alignment Scores:
AR271701 876 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 4 from patent US 6503744.

```

```

Pred. No.: 1.91e-185 Length: 876
Score: 1528.00 Matches: 285
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x BD249792 (1-876)

Qy      1  MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      1  ATGAAAAAGTATTTATTTGCTGGAAATGCACCAAGTTTAAAAAGAAATTCGATTATTCAAGG 60
Qy      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      61  CTACCAAAAGATTTTGTGATATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy      41  LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      121  CTTCGTAATAAATTCAAAGCAGTATTTTACAAATCTCTGGTCTTTTTCGAACAATACTAC 180
Qy      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      181  ACTTTAAACATTTAAATCCAAATCAAGAATATGAGCCGAACATAATATGTGTTCTAAT 240
Qy      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      241  TACAAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAACCTTTTACGATTATTTTCC 300
Qy      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      301  GATGCTCATTTGGGATATGATTTTAAACAACCTTAAAGAATTTAATGCTTATTTTAAA 360
Qy      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      361  TTTCCAGAAATTTATCTCAATCAAAGAAATTAACCTCAGGAGTCTATATGTGTGCAGTAGCT 420
Qy      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      421  ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTCATTTTATCAAAATGGGTCA 480
Qy      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db      481  TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAACCTGGCTCTCGATTTTAAAAAT 540
Qy      181  AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      541  GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Qy      201  GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuLeuAlaAsnPhe 220
Db      601  GAAAAAATTTACAAAATAAAACTATATGCTTATGCTCTAACAGTCTTTTAGCAAAATTTT 660
Qy      221  IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db      661  ATAGAACTAGCGCCAAATTTAAATTTTATCATACAAAGAAAAATTAACACTACTACT 720
Qy      241  LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      721  AAAGATATATCTATACCTCTTAGTGAGCTTATCGAAATTTTCAAAAAATATTAATTTT 780
Qy      261  LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db      781  AAAAAATAAAATTAAGAAAAATATTTATTAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy      281  ProSerAspIleLysHisTyrPheLysGlyLys 291
Db      841  CCTAGTATATAAGCAATATTTTCAAAGGAAAA 873

RESULT 29
AR271701
LOCUS
DEFINITION Sequence 4 from patent US 6503744.

```

```
ACCESSION AR271701
VERSION AR271701.1 GI:29703246
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6503744-A 4 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
    Location/Qualifiers
    source 1..876
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores: 1.91e-185 Length: 876
Pred. No.: 1528.00 Matches: 285
Score: 1528.00
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 2 Gaps: 0
US-10-734-719-9 (1-291) x AR271701 (1-876)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAATAATGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGTATGATTTAGATGCAATCAATTTTATTGGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATTCAAAGCAGTATTACAACTCTGGTCTTTTGAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATAATTAATGTTCTTAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTGTAAAAACCTTTTACGATTATTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAATTTAATGCTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACGAAATTTATCTCAATCAAGAATTAACCTCAGGAGTCTATATGTCAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTCTGGAATGATTTTATCAAAATCGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAACTGGCTCTCTGAATTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCCGACATAGTAAAAATACAGATATATAAGAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTCAAAATAAATACTATATTGCTTATGCTCTTACAGCTCTTTAGCAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
DB 661 ATAGAACTAGCGCAAAATTTAAATTTTATCATACAGAAAAAATAACTACTACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATATCTCATACCTCTTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 780
QY 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAAATAAATAATTAAGAAATATTTATTAAGTTGATAAAGATCTTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAGCATTTTTCAAAGGAAAA 873
RESULT 30
AR481783
LOCUS AR481783
DEFINITION Sequence 4 from patent US 6699705.
ACCESSION AR481783
VERSION AR481783.1 GI:47243418
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
JOURNAL Patent: US 6699705-A 4 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
    Location/Qualifiers
    source 1..876
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores: 1.91e-185 Length: 876
Pred. No.: 1528.00 Matches: 285
Score: 1528.00
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 2 Gaps: 0
US-10-734-719-9 (1-291) x AR481783 (1-876)
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAATAATGATTATTCAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
DB 61 CTACCAATGATTTTGTATGATTTAGATGCAATCAATTTTATTGGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATTCAAAGCAGTATTACAACTCTGGTCTTTTGAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACATTTAATCCAAATCAAGATATGAGCCGAATAATTAATGTTCTTAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TACAACCAAGCTCATCTAGAAAAATGAAAATTTGTAAAAACCTTTTACGATTATTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAATTTAATGCTTATTTAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTACGAAATTTATCTCAATCAAGAATTAACCTCAGGAGTCTATATGTCAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGATACAAAGAAATTTATCTTCTGGAATGATTTTATCAAAATCGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGATACCAACAGAAATCTTTTAAAACTGGCTCTCTGAATTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATCGCTCACACTATATCCGACATAGTAAAAATACAGATATATAAGAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAATTTCAAAATAAATACTATATTGCTTATGCTCTTACAGCTCTTTAGCAATTTT 660
```



```

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGCAAAATTTATCTCAATCAAGAATTTACCTCAGGAGTCTATATGTGTGAGTAGCT 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTCGAATTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACCAAGAAATCTTTTAAAACTGGCTCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGGAGCTTTATGGAATAATTTTCAAAAAATTTAAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAATTAAGAAATAATTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 31
AR527382 AR527382 876 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 4 from patent US 6723545.
ACCESSION AR527382
VERSION AR527382.1 GI:53914403
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Polypeptides having .beta.-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 4 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
    source
        Location/Qualifiers
            1..876
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.91e-185 Length: 876
Score: 1528.00 Matches: 285
Percent Similarity: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR527382 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTTATGCTGGAATGGACCAAGTTTAAAGAAATTTAGATTATTCAGG 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40

```

```

Db 61 CTACCAATGATTTTGATGTATTTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATTTCAAAGCAGTATTTTACAAATCCTGGTCTTTTTTTTGAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACAATTTAAATCCAAATCAAGAATATGAGACCGAACTAATTTATGTGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAATTTTGTAAAAAATCTTTTACGATTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTTAAACAACCTTAAAGAAATTTAATGCTTTATTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGAAATTTATCTCAATCAAGAATTTACCTCAGGAGTCTATATGTGTCAGTAGCT 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAAAACAAGAAATCTTTTAAAACTGGCTCTGATTTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAACTATATTGCTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAAGAAAAATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTTAGGAGCTTTATGGAATAATTTTCAAAAAATTTAAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAAAATTAAGAAATAATTTTATACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 32
AR609661 AR609661 876 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 4 from patent US 6825019.
ACCESSION AR609661
VERSION AR609661.1 GI:56664961
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 4 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
    source
        Location/Qualifiers
            1..876

```

```
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
  1.91e-185      Length: 876
Pred. No.:      1528.00    Matches: 285
Score:          98.3%      Conservative: 1
Percent Similarity: 97.9%  Mismatches: 5
Best Local Similarity: 97.8% Indels: 0
Query Match: 2          Gaps: 0
DB:

US-10-734-719-9 (1-291) x AR609661 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCTCGAAATGCACCAAGTTTAAACAGAAATTGATTATTCAAGG 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGATGCAATCAATTTTATTTCGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGTGTAATAAATTCAAAGCAGTATTTTACAATCTCTGGTCTTTTGTGAACAATACTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGCCGAACTAATTATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAACTTTTACGATTATTTCCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTCCGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAAGAAATTTATCTCAATCAAGAAATTTACCTCAGGAGTCTATATGTGTCAGTAGCT 420

Qy 141 IleAlaLeuGlyTyrLysGluIleThrSerGlyValTyrMetCysAlaValAla 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTGGAATTTGATTTTATCAAAATGGGTCA 480

Qy 161 SerTyrAlaPheAspThrLysLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGTATACCAACCAAGAAAATCTTTTAAAACTGGCTCCTGATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTCAAAAATAAACTATTTGCTTATGCTTAAACAGTCTTTTATAGCAAAATTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGACTAGCGCCAAATTTAAATTCAAATTTTATCATCAAGAAAAAATAAATCACTACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATCTCATACCTCTTAGTGAGGCTTATCGAAAATTTTCAAAAAATTTAATTTT 780

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAATTAAGAAAAATATTATTATCAAGTTGATAAAGATCTATTAAAGATTA 840

Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGGCATTATTATTTCAAGGAAAA 873
```

```

QY 201 GLLYSLThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAAACATATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTAGTAGCGCTTATGGAATTTTCAAAAAATTAATTAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATAAGAAAAATATTTTATACAAGTTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTAATTTCAAGGAAAA 873

RESULT 34
AR691835
LOCUS AR691835 876 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 4 from patent US 6911337.
ACCESSION AR691835
VERSION AR691835.1 GI:74479546
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAc transferase
JOURNAL Patent: US 6911337-A 4 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.91e-185 Length: 876
Pred. No.: 1528.00 Matches: 285
Score: 98.3% Conservative: 1
Best Local Similarity: 97.9% Mismatches: 5
Query Match: 97.8% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR691835 (1-876)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTATTATTGCTGGAATTTGACCAAGTTTAAAGAAATATGATTATCAAGG 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGTATTAGTAGCAATCAATTTTATTTTGAAGATAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATTCAAAGCAGTATTTTACAATCTCGTCTTTTGAACATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTAAATCCAAAATCAAGAATATGAGACCGCAACTAAATATGTTGTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAGAACTTTTACGATTATTTTCCT 300

```

```

QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTAAAAACAACCTTAAAGAAATTAATGCTTATTTAAAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCACGAAATTTATCTCAATCAAGAAATTAACCTCAGGAGTCTATATGTGTCAGTAGCT 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAATGGGTCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAGAAATCTTTTAAACTGGCTCTGATTTTAAAAAT 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAATAAAACATATATGCTTATGCTTAAACAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTAAATTCAAATTTTATCATACAAGAAAAATACTACACT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTCATACCTCTAGTAGCGCTTATGGAATTTTCAAAAAATTAATTAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATAAGAAAAATATTTTATACAAGTTTGATAAAGATCTATTAAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTAATTTCAAGGAAAA 873

RESULT 35
AR6934425
LOCUS AR6934425 876 bp DNA linear PAT 05-JAN-2004
DEFINITION Sequence 2 from Patent WO0207942..
ACCESSION AR6934425
VERSION AR6934425.1 GI:40641672
KEYWORDS
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
REFERENCE 1
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics
JOURNAL Patent: WO 0207942-A 2 26-SEP-2002;
National Research Council of Canada (CA)
FEATURES
source Location/Qualifiers
1..876
/organism="Campylobacter jejuni"
/mol_type="unassigned DNA"
/db_xref="taxon:197"
1..876
/note="unnamed protein product; bifunctional alpha-2,3/alpha 2,8-sialyltransferase Campylobacter sialyltransferase II (cstII) from C. jejuni strain OH4384 (ORF 7a of lipooligosaccharide (LOS) biosynthesis locus)"
/transl_table=1
/protein_id="CAF04166.1"
/db_xref="GI:40641673"
/translation="MKKVIAGNGPSLKEIDYSLPNDFDVFRCNQFYEDKYLGKK CKAFVNPILFFEQYYTLKHLLIQNQEYETELMCSNYNQAHLENENFVKTFDYDFDA

CDS

```

HLGYDFFKQLKDFNAYKFKHEIYFNQRTITSGVYMCVAIAALGYKEIYLSGIDFYONGS  
SYAFDTQKNLLKLPNFNDNSHYIGHSKNTDIKALEFLEKTYKIKLYCICPNLSLLA  
NFIELAPNLNENFIIQENKNTVDILIPSEAYGKFSKNINFKIKIKENIYVYKLIKID  
LLRLPSDIKHVFKGK"

ORIGIN

Alignment Scores:

Pred. No.: 4.63e-185 Length: 876  
Score: 1525.00 Matches: 283  
Percent Similarity: 98.6% Conservative: 4  
Best Local Similarity: 97.3% Mismatches: 4  
Query Match: 97.6% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AX934425 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTCATTATTCAAGA 60  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTGGTATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACATAT 120  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAATAAATGCAAGCGATATTTACAATCCTATTCTTTTTTGAACAATACCTAC 180  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 181 ACTTTAAACATTTAAATCCAAATCAAGAATATGACCGAACAATAATATGTTCTTAAT 240  
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCC 300  
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTCAAACAACTTTAAAGATTTTAAATGCTTATTTTAAA 360  
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTCAAGAAATTTATTTCATCAAGAATACCTCAGGGTTTATATGTCAGTAGCC 420  
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 480  
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTTANAATTTGGCTCCTTAATTTTAAAAAT 540  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATTAATTCACACTATATCGGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 600  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTACAAAAATAAACTATATTGCTTATGCTTAACAGTCTTTTAGCAAAATTTT 660  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACCTAGCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATACACT 720  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 721 AAAATATACCTCATCACTCTTAGTGAGGCTTAGGAAAAATTTTCAAAAAATTAATTAATTT 780  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 781 AAAAAATAAAAATTAAGAAAAATTTATTATACAAGTTTGATAAAAGATCTATTAAAGATTA 840  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291

Db 841 CCTAGTGATATAAGCAATTTTCAAAGGAAA 873

RESULT 36

LOCUS

DEFINITION

BD249791 876 bp DNA linear PAT 17-JUL-2003  
Campylobacter glycosyltransferases for biosynthesis of  
gangliosides and ganglioside mimics.

ACCESSION

BD249791

VERSION

BD249791.1 GI:33059561

KEYWORDS

JP 2002535992-A/2.

SOURCE

Campylobacter jejuni

ORGANISM

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

REFERENCE

1 (bases 1 to 876)

AUTHORS

Gilbert, M. and Wakarchuk, W.W.

TITLE

Campylobacter glycosyltransferases for biosynthesis of

JOURNAL

gangliosides and ganglioside mimics

COMMENT

Patent: JP 2002535992-A 2 29-OCT-2002;

NATIONAL RESEARCH COUNCIL OF CANADA

OS Campylobacter jejuni

PN JP 2002535992-A/2

PD 29-OCT-2002

PF 01-FEB-2000 JP 2000597438

PR 01-FEB-1999 US 60/118213, 31-JAN-2000 US 06/495406 PI

MICHEL GILBERT, WARREN W WAKARCHUK

PC C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC

Campylobacter glycosyltransferases for biosynthesis of CC

gangliosides and

CC ganglioside mimics

PH Key Location/Qualifiers

FT source 1..876

FT Location/Qualifiers

1..876

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/db\_xref="taxon:197"

ORIGIN

Alignment Scores:

Pred. No.: 4.63e-185 Length: 876

Score: 1525.00 Matches: 283

Percent Similarity: 98.6% Conservative: 4

Best Local Similarity: 97.3% Mismatches: 4

Query Match: 97.6% Indels: 0

DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x BD249791 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20

Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTCATTATTCAAGA 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

Db 61 CTACCAATGATTTTGGTATTTTAGATGTAATCAATTTTATTTTGAAGATAAATACATAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

Db 121 CTGTGTAATAAATGCAAGCGATATTTACAATCCTATTCTTTTTTGAACAATACCTAC 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80

Db 181 ACTTTAAACATTTAAATCCAAATCAAGAATATGACCGAACAATAATATGTTCTTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100

Db 241 TACAACCAAGCTCATCTAGAAATGAAAAATTTTGTAAAACTTTTACGATTAATTTTCC 300

Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

```
Db 301 GATGCTCATTTGGGATATGATTTTTTCAACAACTTTAAAGATTTTAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTATGCTTTTGATACAAACAAATCTTTTAAATTTGCTTCTCTAAATTTTAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTACAAATATAAATCTATATGCTTATGCTTACACAGTCTTTTACCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTTAAATTTTAAATTTTATCATACAGAAATATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCATACCTCTAGTAGGCTTATGGAAATTTTCAAAAAATATTTAAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAATATATTTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 37
AR271700
LOCUS AR271700 876 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 2 from patent US 6503744.
ACCESSION AR271700
VERSION AR271700.1 GI:29703245
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6503744-A 2 07-JAN-2003;
National Research Council of Canada, Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 4.63e-185 Length: 876
Pred. No.: 1525.00 Matches: 283
Score: 98.6% Conservative: 4
Percent Similarity: 97.3% Mismatches: 4
Best Local Similarity: 97.6% Indels: 0
Query Match: 2 Gaps: 0
DB:

US-10-734-719-9 (1-291) x AR271700 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
|||||
```

```
Db 1 ATGAAAAAGCTTATTTATTTGCTGGAATGGACCAAGTTTAAAGAAATGATTTATTTCAAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGAATGATTTTAGATGATTTCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAAATGCAAGGCGAGTATTTTACAATCTTATTTCTTTTGTGAAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACATTTTAAATCCAAAATCAAGAAATATGAGACCGAACTAAATTTATGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACCTTTTTCAGATTTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAAGAAATTTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTTGCGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAAATGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACAAACAAATCTTTTAAATTTGGCTCTCTCTCTCTCTCTCTCT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAAATTCACATATATCGGACATAGTAAAAATACAGATATAAAGCTTTTAAAGATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTTCAAAAAATAAATACTATTTGCTTATGCTTCTTAAACAGTCTTTTACCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAATTTTAAATTTTAAATTTTATCATACAGAAATATAACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCATACCTCTAGTAGGCTTATGGAAATTTTCAAAAAATATTTAAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAAAAATTAAGAAATATTTTATTTACAAAGTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873

RESULT 38
AR481782
LOCUS AR481782 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 2 from patent US 6699705.
ACCESSION AR481782
VERSION AR481782.1 GI:47243417
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 2 02-MAR-2004;
```

National Research Council of Canada; Ottawa;  
CAX;

## FEATURES

Location/Qualifiers

source

1..876

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.63e-185 Length: 876  
Score: 1525.00 Matches: 283  
Percent Similarity: 98.6% Conservative: 4  
Best Local Similarity: 97.3% Mismatches: 4  
Query Match: 97.6% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR481782 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTATTATTGCTGGAAATGACCAAGTTTAAAGAAATGATTATTCAGA 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAAAATGATTTTGTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTGTGTAATAAATGCAAGCGAGTATTTACAACTCTATTCTTTTGAACAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCAAAATCAAGAAATAGACCGAACTAATTAATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATAGAAATGAAATTTTGTAAAGTATTTTACGATTATTTCTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTATTGGGATATGATTTTCAACCAACTTAAAGATTTAAATGCTTATTATTA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysSerAsn 140  
DB 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTGTGAGTAGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTATTTTATCAAAATGGGTCA 480

QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
DB 721 AAAGATATACTATACCTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 780

QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
DB 781 AAAAAATAAAAATTTAAAGAAATATTTATTACAAAGTTGATAAAGATCTATTAAAGATTA 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
DB 841 CCTAGTGATATAAGCATTATTTCAAAGGAAAA 873

RESULT 39  
AR527381  
LOCUS AR527381 876 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 2 from patent US 6723545.  
ACCESSION AR527381  
VERSION AR527381.1 GI:53914402  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 876)  
AUTHORS Gilbert,M. and Warkachuk,W.W.  
TITLE Polypeptides having .beta.-1,4-GalNAC transferase activity  
JOURNAL Patent: US 6723545-A 2 20-APR-2004;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
Location/Qualifiers  
source 1..876  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 4.63e-185 Length: 876  
Score: 1525.00 Matches: 283  
Percent Similarity: 98.6% Conservative: 4  
Best Local Similarity: 97.3% Mismatches: 4  
Query Match: 97.6% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR527381 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 1 ATGAAAAAGTATTATTGCTGGAAATGACCAAGTTTAAAGAAATGATTATTCAGA 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 61 CTACCAAAATGATTTTGTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr 60  
DB 121 CTGTGTAATAAATGCAAGCGAGTATTTACAACTCTATTCTTTTGAACAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 181 ACTTTAAACATTTAATCAAAATCAAGAAATAGACCGAACTAATTTATGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
DB 241 TACAACCAAGCTCATAGAAATGAAATTTTGTAAAGTATTTTACGATTATTTCTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 301 GATGCTATTGGGATATGATTTTCAACCAACTTAAAGATTTAAATGCTTATTATTA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTGTGAGTAGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTATTTTATCAAAATGGGTCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
DB 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAAAATTTGGCTCTCTAATTTTAAAAAT 540

```
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
|||
Db 541 GATAATTCACACATATATCGGACATAGTAAATACAGATATAAAGCTTTAGAAATTTCTA 600
|||
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
|||
Db 601 GAAAAAATTTACAAAATAAACTATATTTGCTTCTTAAACAGTCTTTTAGCAAAATTTT 660
|||
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnLysAsnAsnTyrThr 240
|||
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
|||
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
|||
Db 721 AAAGATATATCTACATCTCTTAGTGAGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780
|||
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
|||
Db 781 AAAAAATAAAATAAAGAAAAATATTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
|||
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
|||
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
|||
RESULT 40
AR609660
LOCUS AR609660 876 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 2 from patent US 6825019.
ACCESSION AR609660
VERSION AR609660.1 GI:566664960
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert M. and Wakarchuk, W. W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 2 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 4.63e-185 Length: 876
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 2 Gaps: 0
US-10-734-719-9 (1-291) x AR609660 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
|||
Db 1 ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAAAATTTGATTATTCAAGA 60
|||
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
|||
Db 61 CTACCAAAATGATTGTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120
|||
Qy 41 LeuGlyLysCysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
|||
Db 121 CTGGTAAAAAATGCAAGGAGATTTTACAAATCTTATCTTTTGTGAACAATACTAC 180
|||
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
|||
Db 181 ACTTTAAAAACATTAAATCCAAATCAAGATATGAGACCGAACTAATATTGTGTTCTAAT 240
|||
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
|||
```

```
Db 241 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAAATTTTACGATTATTTTCCCT 300
|||
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
|||
Db 301 GATGCTCATTTGGGATATGATTTTTCAAACAACTTAAAGATTTTAAATGCTTATTTTAAA 360
|||
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
|||
Db 361 TTTTCAAGAAATTTATTTCAATCAAGAAATTTTCTTCGGAATTTGATTTTATCAAAATGGGTCA 420
|||
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
|||
Db 421 ATAGCCCTAGGATACAAAGAAATTTTATCTTCGGAATTTGATTTTATCAAAATGGGTCA 480
|||
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
|||
Db 481 TCTTATGCTTTTGATACTAAAAAATAAATCTTTTAAAAATTTGGCTCCTCTAAATTTTAAAAAT 540
|||
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
|||
Db 541 GATAATTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTTAGAATTTCTA 600
|||
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
|||
Db 601 GAAAAAATTTACAAAATAAATATATTGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660
|||
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnLysAsnAsnTyrThr 240
|||
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAACTACACT 720
|||
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
|||
Db 721 AAAGATATATCTACATCTCTTAGTGAGCTTATGGAATAATTTTCAAAAAATATTAAATTT 780
|||
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
|||
Db 781 AAAAAATAAAATAAAGAAAAATATTATTACAAGTTGATAAAGATCTATTAAAGATTA 840
|||
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
|||
Db 841 CCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873
|||
RESULT 41
AR689938
LOCUS AR689938 876 bp DNA linear PAT 12-SEP-2005
DEFINITION Sequence 2 from patent US 6905867.
ACCESSION AR689938
VERSION AR689938.1 GI:74471946
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity
JOURNAL Patent: US 6905867-A 2 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 4.63e-185 Length: 876
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 2 Gaps: 0
```

US-10-734-719-9 (1-291) x AR689938 (1-876)

```
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGA 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGTGTAATAAATGCAAGCGAGTATTTTACAATCTATTCTTTTGTGAACAACTACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACACATTTAATCCAAATCAAGAATATGAGACCGAACTTAATTTTGAAGATAAATACTAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACACACCAAGCTCATCTAGAAATGAAATTTTGTGAAATCTTTTACGATTTATTTCTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTTCAGCAATTTATTCAATCAAGAATTTACCTCAGGGGTTTATATGTGTCAGTAGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAAATTGATTTTATCAAAAATGGTGCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACAAAGAAATCTTTTAAATTTGGCTTCTTAATTTTAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTCACATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATCTACAAAATAAACTATATGCTTATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660

QY 241 LysAspIleLeuIleProSerGluAlaTyrGlyLysPheSerLysAsnIleAsn 260
Db 721 AAAGATATACATACATCTTAGTGAGCTTATGGAAATTTTCAAAAATATTAATTTT 780

QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 781 AAAAAATATAAAATTAAGAAATATTTTACAAAGTTGATATAAGATCTATTAAAGATTA 840

QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
```

RESULT 42

AR691834

LOCUS

DEFINITION

Sequence 2 from patent US 6911337.

ACCESSION

AR691834

VERSION

AR691834.1 GI:74479545

KEYWORDS

Unknown.

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 876)

AUTHORS

Gilbert,M. and Wakarchuk,W.W.

TITLE

Nucleic acids encoding .beta.-1,4-GalNAc transferase

JOURNAL

Patent: US 6911337-A 2 28-JUN-2005;

CAX;

FEATURES

Location/Qualifiers

1..876

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 4,63e-185

Score: 1525.00

Percent Similarity: 98.6%

Best Local Similarity: 97.3%

Query Match: 97.6%

DB: 2

US-10-734-719-9 (1-291) x AR691834 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20

Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTTCAAGA 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

Db 61 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

Db 121 CTTGTGTAATAAATGCAAGCGAGTATTTTACAATCTATTCTTTTGTGAACAACTACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80

Db 181 ACTTTAAACACATTTAATCCAAATCAAGAATATGAGACCGAACTTAATTTTGAAGATAAATACTAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100

Db 241 TACACCAAGCTCATCTAGAAATGAAATTTTGTGAAATCTTTTACGATTTATTTCTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

Db 301 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

Db 361 TTTTCAGCAATTTATTCAATCAAGAATTTACCTCAGGGGTTTATATGTGTCAGTAGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGAAATTTGATTTTATCAAAAATGGTGCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180

Db 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAATTTGGCTTCTTAATTTTAAAAT 540

QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

Db 541 GATAATTCACATATATCGGACATAGTAAATAACAGATATATAAGCTTTTAGAATTTCTA 600

QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220

Db 601 GAAAAAATCTACAAAATAAACTATATGCTTATGCTTATGCTTAAACAGTCTTTTAGCAAAATTTT 660

QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240

Db 661 ATAGAACTTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATAAATACTACT 720

QY 241 LysAspIleLeuIleProSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

Db 721 AAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAAATTTTCAAAAAATATTAATTTT 780



```

Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAAAAATAAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATATTATTTCAAGGAAAA 873

RESULT 43
AF216647
LOCUS
DEFINITION
  Campylobacter jejuni alpha-2,3/8-sialyltransferase (cst-II) gene,
  complete cds.
ACCESSION
  AF216647
VERSION
  AF216647.1 GI:17221830
KEYWORDS
  Campylobacter jejuni
SOURCE
  Campylobacter jejuni
  Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
  Campylobacteraceae; Campylobacter.
REFERENCE
  1 (bases 1 to 876)
  Gilbert M., Michniewicz, J. and Wakarchuk, W. W.
  Cloning and sequencing of the cst-II gene from Campylobacter jejuni
  ATCC 700297
  Unpublished
JOURNAL
  2 (bases 1 to 876)
  Gilbert M., Michniewicz, J. and Wakarchuk, W. W.
  Direct Submission
  Submitted (15-DEC-1999) Institute for Biological Sciences, National
  Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
  0R6, Canada
FEATURES
  source
    1..876
    /organism="Campylobacter jejuni"
    /mol_type="genomic DNA"
    /strain="ATCC 700297"
    /serotype="Penner 0:19"
    /db_xref="ATCC:700297"
    /db_xref="taxon:197"
    1..876
    /gene="cst-II"
    /gene="cst-II"
    /gene="cst-II"
    /note="bifunctional sialyltransferase"
    /codon_start=1
    /transl_table=11
    /product="alpha-2,3/8-sialyltransferase"
    /protein_id="AAU36462.1"
    /db_xref="GI:17221831"
    /translation="MKKVIIAGNGPSLKEIDYSLPNDPVDVFRNCNQFYFEDKYVLGKK
    CKAVFYNPILFEOYVYKLHIOQEYETELIMCSNYNQAHLENENFKVTFDYFPDA
    HLGVDYFKQLKDFNAYPKFHEIYNQRTISGVYMCVAIALGYKEIYLSGIDFYONGS
    SYADFQKQLLKLAFNPKNDNSHYIGHSKNTDIKALEFLEKTKIKLYKLCPSNLLA
    NFIELAPLNSNFTLIQKNNYTKDILIPSEAYGKFSKNINFKKIKENIYYKLKID
    LLRLPSDIKHFKGK"

ORIGIN
Alignment Scores:
Pred. No.: 4.63e-185 Length: 876
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 15 Gaps: 0

US-10-734-719-9 (1-291) x AF216647 (1-876)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTATTCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

```

```

Db 61 CTACCAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAACATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGGCAGTATTTTACAACTCCTATTCTTTTTCGAAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACAATTTAATCCAAATCAAGAATATGAGACCGAACTAATTAATGTTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TACAACCAAGCTCATCTAGAAAATTTGTAATAAATCTTTTACGATTATTTTCTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTACAGAAATTTATTTCAATCAAGAATTTACTCAGGGGCTCTATAIGTGTGAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAAAATTCGCTCCTAAATTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTCACACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTTACAAAAATAAACTATATTCCTTATTCCTTAAACAGTCTTTTAGCAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAAGAAAAAATACTACACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATACTACATACCTTCTAGTGAGGCTTATGGAAAAATTTCAAAAAATATTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAAATAAATAAATAAAGAAATATTTATTAAGATTTGATAAAGATCTATTAAAGATTA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTGATATAAGCAATTTATTTCAAGGAAAA 873

RESULT 44
AY644679
LOCUS
DEFINITION
  Campylobacter jejuni strain CF 93-6 LOS outer core biosynthesis
  gene cluster, partial sequence.
ACCESSION
  AY644679
VERSION
  AY644679.1 GI:55420298
KEYWORDS
  Campylobacter jejuni
  Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
  Campylobacteraceae; Campylobacter.
SOURCE
  Campylobacter jejuni
  Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacter.
  1 (bases 1 to 6047)
REFERENCE
  Koga, M., Gilbert M., Li, J., Koike, S., Takahashi, M., Furukawa, K.,
  Hirata, K. and Yuki, N.
  Antecedent infections in Fisher Syndrome: a common pathogenesis of
  molecular mimicry
  Unpublished
JOURNAL

```



```

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheIeu 200
Db 5411 GATAAATTCACACTATATCGGACATAGTAAATAACAGATATAAAAGCTTTAGAAATTTCTA 5470
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 5471 GAAAAAATTACAAAATAAACTATATGCTTATGCTCTAACAGTCCTTTAGCAAATTTT 5530
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIlelleGlnGluLysAsnAsnTyrThr 240
Db 5531 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAAATAAATCACTACT 5590
Qy 241 LysAsnIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 5591 AAAGATATACTCATACCTCTAGGAGCTTATGGAATAATTTTCAAAAAATATTAAATTTT 5650
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 5651 AAAAAATAAAATAAAGAAATATTATTACAGTTGATAAAGATCTATTAGATTA 5710
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 5711 CCTAGTATATAAGCATTTATTTCAAAGGAAAA 5743

RESULT 45
AF400048 11442 bp DNA linear BCT 03-JAN-2002
LOCUS
DEFINITION
Campylobacter jejuni strain ATCC 43438 heptosyltransferase I (waac)
gene, partial cds; lipid A biosynthesis acyltransferase (htrB),
putative two-domain glycosyltransferase, putative
glycosyltransferase, beta-1,4-N-acetyl-galactosaminyltransferase
(cgtA), beta-1,3-galactosyltransferase (cgtB),
alpha-2,3-alpha-2,8-galactyltransferase (cstII), putative stialic
acid synthase (neuB), putative N-acetylglucosamine-6-phosphate
2-epimerase (neuC), CMP-Neu5Ac synthetase (neuA), putative
acyltransferase, and putative glycosyltransferase genes, complete
cds; and heptosyltransferase II (waar) gene, partial cds.
AF400048
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 11442)
Gilbert, M., Karwaski, M.F., Bernatchez, S., Young, N.M., Taboada, E.,
Michniewicz, J., Cunningham, A.M. and Wakarchuk, W.W.
The Genetic Bases for the Variation in the Lipo-oligosaccharide of
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF
STALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
J. Biol. Chem. 277 (1), 327-337 (2002)
11689567
JOURNAL
PUBMED
2 (bases 1 to 11442)
Gilbert, M., Michniewicz, J., Karwaski, M.-F., Cunningham, A. and
Wakarchuk, W.W.
Direct Submission
Submitted (15-JUL-2001) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
source
1. .11442
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43438"
/db_xref="ATCC:43438"
/note="taxon:197"
<1. .348
/gene="waac"
<1. .348
/gene="waac"
/codon_start=1
gene
CDS
complement (5043..5936)
/gene="cgtB"
complement (5043..5936)

```

```

/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAK91719.1"
/db_xref="GI:15193215"
/translation="FAKEVLNLSGIDETHIELAPKFSLEELMAFTKMDLLIIGNDSPG
THLAFALNKASITIFGATPSYRNAPQTHINKIIDAGKIQNAKHIDKSDFCITRISEE
DIFKLAKGLNEK"
338..1225
/gene="htrB"
338..1225
/gene="htrB"
/codon_start=1
/transl_table=11
/product="lipid A biosynthesis acyltransferase"
/protein_id="AAK91720.1"
/db_xref="GI:15193216"
/translation="MKNSDRIYLSLYIYLKFFVTMPPECILHFLALIVARITFYLNKK
HRKIDINLQICFPQYTKERDKLSLKIYENFAQFGIDCLONQNTTKERILANKVFN
ENLIDALALKRPITFTAHYGNWEILSLAVAAYKCAISIVGKILKSEIMYELISQNR
TQDIELDKKGGIROMLSALKGRALGILTDODCVKESVRLKFNKEVNYQMGASL
IAQRNALIIPVAYKEDGKFCIEFFVKYKDSQNASLEELTLVQAQSCSEMIKRPWEY
FFHRRFASYNKEIYKDAK"
1222..2769
/note="orf3"
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAK91721.1"
/db_xref="GI:15193217"
/translation="MNLKQISAIIVKNAKQTLLECLNSLKDDEIILLNESSDNTL
NIANEFKDFANLHYHSAFGPGALKNLALSAYAKNDWILSIDADEVLENEGICKELN
LEQEDNIIALSRNLKYGEWIKACGMPDYVVRIFNKNFTFNDNLVHESLVPSNA
KILYKNGRIYAFRDISHLIDKQYYSLSLWAKONHKSSGVLKANLRAFMPFRNYF
LKNGLGYGKFIISVCNALGTFFKIMKYLQKQKPKCALIITTYNOKERLUKLVLD
SNQNLAFNLVLIADGSKEDTARLIEYQKDFCPKLHIWQEGSESEILKNNYSL
KASKEVYIILIDGMILBKDFVQNHLEPAKRVILQGSRVILNQSESEILKNNYSL
APKNGKYNORSIDFLAKCVYKFSKTKTKFKKSVLQVSKTCNMGYKQDFEAEQFN
EKFTMGREDSEFFVARFLNDGVFKLFNALAYHIYHEENSKNMLEINHQIYLEIK
NEKTTWR"
2773..3957
/note="orf4"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAK91722.1"
/db_xref="GI:15193218"
/translation="MKYGVWPIYVNVKYLKECLDSVINQTYTNLEIILVNDGSTDE
NSLNTAKYTLKDKRITLFDKNGGLSSARNVGIYFSGEYKLNKQTHIKENSLIEF
QLDGNPNYIYKAYKSSOAFNEKDLTNLTYPIDYIIFLDSNWKLCIIEBCVIRM
KNVDLWFDHCTYEDNTKNHKHTRMEIFDFKSCILTPKEYANRALSVGSRDISFG
WNGVIDFNFLQIKLFIINLINEDIHFGIILFASANKIYVLSQKLYLCRLRANSIN
HDKKITKANVSEYFKDIYETFGENAKENYKLAASRVITALKLIEFFKQDKSENALA
IKETPLPCYAKKALMKLKKFKDPLNLKEQLVLKPKFIQTKLPYDIWFKQKIKYLI
KI"
4013..5071
/gene="cgtA"
4013..5071
/gene="cgtA"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetyl-galactosaminyltransferase"
/protein_id="AAK91723.1"
/db_xref="GI:15193219"
/translation="MLFQSYFVKIICLFIPFRKIRHKIKKTKLLKNIQDKIDSYLPK
KTLIQINKYNNEDLILKNAKIIIGGHGKGVFNDEKSKDPKPLNPAWAFIRKNEAITL
KASLEIILPAIQRGVIGVNDCTDGSSEIILEFCOYPSPIKYPYEIQIKNPSEEN
KLYSYNYVASFIPKDEWLLIKIDVDHYDAKLYKSFYIPRKNYHVISYRIDFIENE
EKFYVYRKEGELKAPGDCAIQNTLNFKEILLEDDTFKWTAKNNIENAKSYEIL
KVRNRYFTTELNNYHFPFIKNYKNDYKQLNWNVLDDDFIKNYKEKLKNOIDPFMLEY
KTLKKVYKLTSSASDKI"
complement (5043..5936)
/gene="cgtB"
complement (5043..5936)

```

```

/gene="cgtB"
/codon_start=1
/transl_table=1
/product="beta-1,3-galactosyltransferase"
/protein_id="AAK91724.1"
/db_xref="GI:15193220"
/translacion="MFKISIIILPTYNVEQYIARIAESICINQTFKNIIEIIVVDDCGSDK
SIDIVKEYAKDDRIKIHNENELKLLRARVEGVKANSPIVIFLDPDDYLELNACEE
CMKILKNEIDLLPFNAFLNNKKIERKLNFOBKCVVKDFELKELLTKNLFWTWA
KVKEYLKAUGLISLENAKINMAEDVLLYFNLINISNTIFHLSKNLYNQINFSI
TKTLTQNIKNTIQEONVLYLLKMQYNFNFTLLKLIBFYLLIEKYSUSSRRNVL
CFKINIFPKIQKFYRLKM"
/gene="cstII"
/6025..6900
/gene="cstII"
/6025..6900
/codon_start=1
/transl_table=1
/product="alpha-2,3-/alpha-2,8-sialyltransferase"
/protein_id="AAK91725.1"
/db_xref="GI:15193221"
/translacion="MKKVIISGNGPSLKEIDYSRLPNDPVRQNFVFEDKYLYGKK
FKAVYNGPFFEQYIYTKHLIQOEYETELIMCSNYNQAHLENENFKTYDFPDA
HLYDFKQKFEWAFKFEHIEYILNQITSGVMCAVAIALGYKEIYLYCLDFQNGS
SYAFDTKQENLLKLPDFKNDRSHVIGHSKNTDIKALEFLEKTKIKLYCLCPNSLLA
NFIELANLNSNFIIOEKNNYTKDILIPSSAYGKFSKNINFKKIKIKENIYYKLKD
LLRLPSDIKHYFKGK"
/6901..7941
/gene="neuB1"
/6901..7941
/gene="neuB1"
/codon_start=1
/transl_table=1
/product="putative sialic acid synthase"
/protein_id="AAK91726.1"
/db_xref="GI:15193222"
/translacion="MKELIKQNIISBEKAPLVVPEIGINHGSLLELAKIMVDAFSA
GAKIKRQTHIVEDMSKAARKVIPGNAKISYEIMQCALDYKDELALKYTEKIGL
VYLSPPSRGALNDMGVSAPFKIGECNENYPLIKHIAAFKPMIVSTGMSTIESI
KPTVKILLDNEIPVLWNTNLYPTPHNLVNLAMLEIKKFEFCMVGSLDTTNLAC
LQAVLGCALVERHFTDSMHRSGPDIVCSMDTKALKELIIQSEQMAIIRGNESKAA
KQEVTTIDFAFASVVSUKDIKKGEVLSMDNIWVKRPGLGGISAAEFENILGKKALRDI
ENDQLSYEDFA"
/7938..9056
/gene="neuC1"
/7938..9056
/gene="neuC1"
/codon_start=1
/transl_table=1
/product="putative N-acetylglucosamine-6-phosphate
2-epimerase"
/protein_id="AAK91727.1"
/db_xref="GI:15193223"
/translacion="MKKILFITGRADYSKIKSLMYRVQNSSEFELYIFATGMHLSKN
FGTVKELYNKFNKIYEFINDYIQDKALATIIDGFSRIANELKPEDLLVWGDRI
EPALAAITGALNLLVAHIEGGEISGTIDDSLRHAI SKLAHILVNDSEFAKRLMQLG
EDKSIPIIGSPDLLELNDKISLSKAKYYDINYNVALLMFHPVTTETISIKNQAD
NLVKALQSKNRYIIVPNNDLGPFLILQSYEEFKNNPRFKLPSLRPEYFITLLKNA
DLFIIGNSCILLKEALYTKTAGILYGSBQNGRLGNENTLKVNANSDEILKAINTHKKQ
DLFSAKLEILDSSKLFLEYLQSGDFFKLSTQKVPKDK"
/9053..9718
/gene="neuA1"
/9053..9718
/gene="neuA1"
/codon_start=1
/transl_table=1
/product="CMP-Neu5Ac synthetase"
/protein_id="AAK91728.1"
/db_xref="GI:15193224"
/translacion="PMSLAIIPARGSGKIGKNKLVLLNNKPLIYYTIKAALNAKSISK
VWVSSDSDEILNYAKSQNVLDIKRPISLAQDDTTSDKVLHLLALFKFYKDYEDVFLQPT
SPUKTNTHNEAFNLKNSNANALISVSECNDKILKAFVNCNDGDLGICNDIYFPMP
RQKLPKTYMNGAIYILKIEFKLNNPFLQSKTKHFLMDESSSIDIDCLDLKKVEQI

```

```

MKK"
/9706..9900
/Note="orf11; premature translational stop codon due to
variation in homopolymeric G-tract"
/codon_start=1
/transl_table=1
/product="putative acetyltransferase"
/protein_id="AAK91729.1"
.
Alignment Scores:
Pred. No.:      Length:      11442
Score:          1525.00      284
Percent Similarity: 98.3%      Conservative: 2
Best Local Similarity: 97.6%      Mismatches: 5
Query Match:      97.6%      Indels: 0
DB:               15          Gaps: 0
US-10-734-719-9 (1-291) x AF400048 (1-11442)
QY      1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      6025 ATGAAAAAGTTATTATTCTCGAAATGGACCAGTTTAAAGAAATGATTATTCAAGG 6084
QY      21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db      6085 CTACCAATGATTTTGATGTAATTAGATCAATCAATTTTATTTTGAAGATAAATACTAT 6144
QY      41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      6145 CTTGGTAAAAAATTCAAAGCAGTATTTTACAATCTCTGGTCTTTTGTGAACAACTACTAC 6204
QY      61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db      6205 ACTTTAAACATTTAATCAAAATCAAGAAATATGAGACCGCACTAATATTATGTGTTCTAAT 6264
QY      81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      6265 TACACCAAGCTCACTAGAAATGAAATTTTGTAAACACTTTTACGATTATTTTCT 6324
QY      101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      6325 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAA 6384
QY      121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      6385 TTTTCAGAAATTTATCTCAATCAAGAAATTTACCTCAGGAGTCTATATGTGTGAGTAGCT 6444
QY      141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      6445 ATAGCCCTAGGATACAAAGAAATTTATCTTTCTGGAAATGATTTTATCAAAATGGGTCA 6504
QY      161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db      6505 TCTTATGCTTTTGATACCAACAAAGAAATCTTTTAAACCTGGCTCTGATTTTAAAAAT 6564
QY      181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db      6565 GATCGCTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 6624
QY      201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db      6625 GAAAAAACTTACAAAATAAAACTATATTGCTTATGCTTACAGCTCTTTAGCAAAATTTT 6684
QY      221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAspTyrThr 240
Db      6685 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACACT 6744
QY      241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db      6745 AAAGATATACTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 6804
QY      261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280

```

```
Db      6805 AAAAAAATAAAATTAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 6864
Qy      281 ProSerAspIleLysHisTyrPheLysGlyLys 291
      |||
Db      6865 CCTAGTGATATAAAGCATTATTTCACAAAGGAAAA 6897

RESULT 46
LOCUS   AX934424
DEFINITION Sequence 1 from Patent WO02074942.
ACCESSION AX934424
VERSION   AX934424.1 GI:40641671
KEYWORDS
SOURCE   Campylobacter jejuni
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
          Campylobacteraceae; Campylobacter.
REFERENCE
AUTHORS Gilbert M. and Wakarchuk W.W.
TITLE    Campylobacter glycosyltransferases for biosynthesis of gangliosides
          and ganglioside mimics
JOURNAL Patent: WO 02074942-A 1 26-SEP-2002;
          National Research Council of Canada (CA)
FEATURES
          Location/Qualifiers
            source
              1..11474
                /organism="Campylobacter jejuni"
                /mol_type="unassigned DNA"
                /db_xref="taxon:197"
                /notes="11.5 kb PCR product from C. jejuni OH4384 including
                LOS biosynthesis locus"

ORIGIN
Alignment Scores:
Pred. No.: 5.55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AX934424 (1-11474)
Qy      1 MetLysLysValIleIleAlaGlyAsnGlyProSerLysLysGluIleAspTyrSerArg 20
Db      6048 ATGAAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAGAAATGGATTATTCAGA 6107
Qy      21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      6108 CTACCAATGATTTTGATGATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
Qy      41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      6168 CTTGGTAAAAAATCAAGGCGAGTATTTTACAATCCTATCTTTTGTGTTTGAACAATACTAC 6227
Qy      61 ThrLysLysHisLeuIleGlnAsnGlnGluTyrGlnThrGluLeuIleMetCysSerAsn 80
Db      6228 ACTTTAAACATTTTAATCCAAAATCAAGAAATATGAGACCGCACTAATATGTTCTAAT 6287
Qy      81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      6288 TACAACCAAGCTATCTAGAAAATGAAAATTTTGTAATAACTTTTACGATTATTTTCCT 6347
Qy      101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      6348 GATGCTCATTTGGGATGATGTTTTTCAACAACCTTAAAGATTTTAAATGCTTTATTTTAAA 6407
Qy      121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      6408 TTTTCAAGAAATTTTATTTCAATCAAGAAATTAACCTCAGGGGCTATATGTTGCGAGTAGCC 6467
Qy      141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCA 6527
```

```
Qy      161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
      |||
Db      6528 TCCTATGCTTTTGATACTAAACAAAAAATCTTTTAAATGGCTCCTTAATTTAAAAAT 6587
Qy      181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
      |||
Db      6588 GATAATTACACTATATCGGACATAGTAAATAATACAGATATATAAAGCTTTAGAAATTTCTA 6647
Qy      201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
      |||
Db      6648 GAAAAAACTTACAAAAATAAACTATATGCTTATGCTTAAACAGCTTTTACCAAAATTTT 6707
Qy      221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
      |||
Db      6708 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACACT 6767
Qy      241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
      |||
Db      6768 AAAGATATACTCATACCTTCTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAATTTT 6827
Qy      261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu 280
      |||
Db      6828 AAAAAAATAAAATTAAGAAAAATATTATTACAAAGTTGATAAAGATCTATTAAAGATTA 6887
Qy      281 ProSerAspIleLysHisTyrPheLysGlyLys 291
      |||
Db      6888 CCTAGTGATATAAGCATTATTTCACAAAGAAAA 6920

RESULT 47
LOCUS   BD249790
DEFINITION Campylobacter glycosyltransferases for biosynthesis of
          gangliosides and ganglioside mimics.
ACCESSION BD249790
VERSION   BD249790.1 GI:33059560
KEYWORDS JP 2002535992-A/1.
SOURCE   Campylobacter jejuni
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
          Campylobacteraceae; Campylobacter.
REFERENCE
AUTHORS Gilbert M. and Wakarchuk, W.W.
TITLE    Campylobacter glycosyltransferases for biosynthesis of
          gangliosides and ganglioside mimics
JOURNAL Patent: JP 2002535992-A 1 29-OCT-2002;
          NATIONAL RESEARCH COUNCIL OF CANADA
COMMENT   OS Campylobacter jejuni
          PN JP 2002535992-A/1
          PD 29-OCT-2002
          PF 01-FEB-2000 JP 2000597438
          PR 01-FEB-1999 US 60/118213,31-JAN-2000 US 06/495406 PI
          MICHEL GILBERT,WARREN W WAKARCHUK
          PC C12N15/09,C12N1/21,C12N9/10,C12N9/88,C12N15/00 CC
          Campylobacter glycosyltransferases for biosynthesis of CC
          gangliosides and
          CC ganglioside mimics
          FH Key
          FT source
            1..11474
              Location/Qualifiers
                /organism='Campylobacter jejuni'
                /mol_type='genomic DNA'
                /db_xref='taxon:197'

ORIGIN
Alignment Scores:
Pred. No.: 5.55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
```

DB:	2	Gaps:	0
US-10-734-719-9 (1-291) x BD249790 (1-11474)			
QY	1	MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
DB	6048	ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAGAAATTTGATTATTTCAAGA	6107
QY	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
DB	6108	CTACCAATGATTTTGTATTTAGATGTATCAATTTTATTTTGAAGATAAATACTAT	6167
QY	41	LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr	60
DB	6168	CTTGCTAAAAATGCAAGCAGTATTTTACAATCTTATTTTGAACAATACTAC	6227
QY	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
DB	6228	ACTTTAAACATTTTAAATCCAAATATGAGACCAAGTATTTTATGTTCTTAAT	6287
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro	100
DB	6288	TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT	6347
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
DB	6348	GATGCTCAITTTGGATATGATTTTTCAAAACAACTTAAAGATTTTATGCTTATTTTAAA	6407
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
DB	6408	TTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTGAGTAGCC	6467
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
DB	6468	ATAGCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA	6527
QY	161	SerTyrAlaPheAspThrLysGlnAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
DB	6528	TCCTATGCTTTTGATACTAAACAAAAAATCTTTTAAAAATTTGGCTCCTCAATTTTAAAAAT	6587
QY	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
DB	6588	GATAATTTCACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTAGAAATTTCTA	6647
QY	201	GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe	220
DB	6648	GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTACAGTCTTTTAGCAAAATTTT	6707
QY	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlyLysAsnAsnTyrThr	240
DB	6708	ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACT	6767
QY	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260

Unclassified.  
1 (bases 1 to 11474)  
Gilbert M. and Wakarchuk, W.W.  
Campylobacter glycosyltransferases for biosynthesis of gangliosides  
and ganglioside mimics  
Patent: US 6503744-A 1 07-JAN-2003;  
National Research Council of Canada; Ottawa;  
CAX;

FEATURES  
source  
1..11474  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 5,55e-184 Length: 11474  
Score: 1525.00 Matches: 283  
Percent Similarity: 98.6% Conservative: 4  
Best Local Similarity: 97.3% Mismatches: 4  
Query Match: 97.6% Indels: 0  
DB: 2 Gaps: 0  
US-10-734-719-9 (1-291) x AR271699 (1-11474)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
DB 6048 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAGAAATTTGATTATTTCAAGA 6107  
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
DB 6108 CTACCAATGATTTTGTATTTAGATGTATCAATTTTATTTTGAAGATAAATACTAT 6167  
QY 41 LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr 60  
DB 6168 CTTGCTAAAAATGCAAGCAGTATTTTACAATCTTATTTTGAACAATACTAC 6227  
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
DB 6228 ACTTTAAACATTTTAAATCCAAATATGAGACCAAGTATTTTATGTTCTTAAT 6287  
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro 100  
DB 6288 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTTCT 6347  
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
DB 6348 GATGCTCAITTTGGATATGATTTTTCAAAACAACTTAAAGATTTTATGCTTATTTTAAA 6407  
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
DB 6408 TTTTCAGCAATTTTATTTCAATCAAGAAATTTACCTCAGGGGCTATATGTGTGAGTAGCC 6467  
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
DB 6468 ATAGCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTCA 6527  
QY 161 SerTyrAlaPheAspThrLysGlnAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
DB 6528 TCCTATGCTTTTGATACTAAACAAAAAATCTTTTAAAAATTTGGCTCCTCAATTTTAAAAAT 6587  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
DB 6588 GATAATTTCACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTTAGAAATTTCTA 6647  
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuProAsnSerLeuLeuAlaAsnPhe 220  
DB 6648 GAAAAAATTTTACAAAATAAACTATATGCTTATGCTTACAGTCTTTTAGCAAAATTTT 6707  
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlyLysAsnAsnTyrThr 240  
DB 6708 ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACT 6767  
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

LOCUS AR271699 11474 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 1 from patent US 6503744.  
ACCESSION AR271699  
VERSION AR271699.1 GI:29703244  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

```
Db 6768 AAGATATACATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATATTAAATTT 6827
Qy 261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu 280
Db 6828 AAAAAATAAAAATTAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 6887
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCATTTATTCAAAGGAAAA 6920

RESULT 49
AR481781
LOCUS AR481781 11474 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 1 from patent US 6699705.
ACCESSION AR481781
VERSION AR481781.1 GI:47243416
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert, M. and Wakarchuk, W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 1 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1. .11474
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 5,55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 0
Query Match: 97.6% Indels: 0
Gaps: 2

US-10-734-719-9 (1-291) x AR481781 (1-11474)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLysAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATTGCTGGAATGCACCAAGTTTAAAGAAATTTGATTATTCAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAATGATTGATGTTATGATGTTATGATGTTATGATGTTATGATGTTATGATGTTAT 6167
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6168 CTGTGTAATAAATGCAAGCGAGTATTTCACATCTATTCTTTTGTGACAACTACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db 6228 ACTTTAAACATTTAATCCAAAATCAAGAATATGAGACCGAATTAATTATGTTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACCTTTTACGATTATTTCCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCAATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTTATTTTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTACGAAATTTATTTCAAATCAAGAATTACCTCAGGGGCTATATGTTGTCAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
```

```
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACATAAAACAAAAATCTTTTAAAAATTCGCTCCTAAATTTAAAAAT 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGACATAGTAAATAACAGATATATAAGCTTTAGAAATTTCTA 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAACTTACAAAAATAAACTATATTGCTTATGCTCTAACAGTCTTTTAGCAATTTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATCAAGAAAAATAACTACACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATATCATATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTATTT 6827
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAATAAAAATTAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTAAAGATTA 6887
Qy 281 ProSerAspLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCATTTATTCAAAGGAAAA 6920

RESULT 50
AR527380
LOCUS AR527380 11474 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6723545.
ACCESSION AR527380
VERSION AR527380.1 GI:53914401
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert, M. and Wakarchuk, W.
TITLE Polypeptides having beta-1,4-GalNAc transferase activity
JOURNAL Patent: US 6723545-A 1 20-APR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1. .11474
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 5,55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
Gaps: 2

US-10-734-719-9 (1-291) x AR527380 (1-11474)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLysAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATTGCTGGAATGCACCAAGTTTAAAGAAATTTGATTATTCAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAATGATTGATGTTATGATGTTATGATGTTATGATGTTATGATGTTATGATGTTAT 6167
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
```

```
Db 6168 CTTGGTAAATAATGCAAGCGAGTATTTTACAAATCCTATTCTTTTTTTTGAACAATACTACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db 6228 ACTTTTAAACATTTTAATCAAATCAAGATATGAGACCGAATACTAATATATGTTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTGAAAACCTTTTACGATTATTTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 CATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGCTATATGTGTGAGTAGGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTC 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGGTC 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGGCATAGTAAATAATACAGATATATAAGCTTTTAGAATTTCT 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTTACAAATTAAGCTATTTTGTCTTATGCTTACAGCTTTTACCAATTTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAATACTACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATACCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 6827
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 6828 AAAAAAATAAAAAATTAAGAAATAATTTATTTACAAAGTTGATAAGAGTCTATTAAAGATTA 6887
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCATTTATTTTCAAGGAAAA 6920
```

```
RESULT 51
LOCUS AR609659 11474 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 1 from patent US 6825019.
ACCESSION AR609659
VERSION AR609659.1 GI:56664959
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
```

```
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert, M. and Wakarchuk, W. W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 1 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
```

```
FEATURES
source
1..11474
/organism="unknown"
/mol_type="genomic DNA"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 5.55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: Gaps: 2

US-10-734-719-9 (1-291) x AR609659 (1-11474)

Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTTTATTATGCTGGAATGGACCAGTTTAAAGAAAAATGATTATTCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6108 CTACCAAAATGATTTTGATGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6167
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr 60
Db 6168 CTTGGTAAAAAATGCAAGGCAGTATTTTACAAATCCTATTCTTTTGTGAACAATACTAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80
Db 6228 ACTTTTAAAAATTTAATCCAAAAATCAAGAAATATGAGACCGAATAATTTATGTGTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAAAATCTTTTACGATTATTTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTACGAAATTTATTTCAATCAAGAATTTACCTCAGGGGCTATATGTGTGAGTAGGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTC 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACAAAGAAATTTTAAATTTTCAAAATTTTAAATTTTAAAAAT 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAATTCACACTATATCGGCATAGTAAATAATACAGATATATAAGCTTTTAGAATTTCT 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATTTTACAAATTAAGCTATTTTGTCTTATGCTTACAGCTTTTACCAATTTT 6707
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAATAATACTACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAAGATATACCTATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATTTAATTT 6827
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuArgLeu 280
Db 6828 AAAAAAATAAAAAATTAAGAAATAATTTATTTACAAAGTTGATAAGAGTCTATTAAAGATTA 6887
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTGATATAAGCATTTATTTTCAAGGAAAA 6920

RESULT 52
LOCUS AR689937 11474 bp DNA linear PAT 12-SEP-2005
DEFINITION Sequence 1 from patent US 6905867.
```



```
ACCESSION AR689937
VERSION AR689937.1 GI:74471945
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding: polypeptides with .beta.1.1,3-galactosyl
JOURNAL transferase activity
Patent: US 6905867-A 1 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source
1. .11474
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 5.55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 2 Gaps: 0
US-10-734-719-9 (1-291) x AR689937 (1-11474)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 6108 CTACCAATGATTTTGTAGTATTAATCAATTTTATTTGAAGATAAATACTAT 6167
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
Db 6168 CTTGGTAAAAAGTTCAGGCGAGTATTTTACAATCCTATCTTTTGTGAACAATAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAAGTAAATATGTTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCAAGAAATTTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTGCGAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACATAACAAAAATCTTTTAAAAATTTGGCTCTCTAAATTTAAAAAT 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAAATTCACACTATATCGGCATAGTAAAAAATACAGATATAAAGCTTTTAGAATTTCTA 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTACAGTCITTTTAGCAAAATTT 6707
```

```
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6708 ATAGACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTACACT 6767
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6768 AAGATATACTCATACCTTCTAGTGAGGCTTATGGAATTTTCAAAAAATTTAATATTT 6827
Qy 261 LysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAATATAAAATTTAAAGAAATATTTATACAAAGTTGATAAAAGATCTATTAGATT 6887
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTATATAAGCATTATTTCAAAGGAAAA 6920
RESULT 53
AR691833
LOCUS AR691833 11474 bp DNA linear PAT 13-SEP-2005
DEFINITION Sequence 1 from patent US 6911337.
ACCESSION AR691833
VERSION AR691833.1 GI:74479544
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding .beta.-1,4-GalNAC transferase
JOURNAL Patent: US 6911337-A 1 28-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source
1. .11474
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 5.55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 2 Gaps: 0
US-10-734-719-9 (1-291) x AR691833 (1-11474)
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6048 ATGAAAAAGTTATTATGCTGGAAATGGACCAAGTTTAAAGAAATGATTATCAAGA 6107
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 6108 CTACCAATGATTTTGTAGTATTAATCAATTTTATTTGAAGATAAATACTAT 6167
Qy 41 LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
Db 6168 CTTGGTAAAAAGTTCAGGCGAGTATTTTACAATCCTATCTTTTGTGAACAATAC 6227
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6228 ACTTTAAACATTTTAAATCCAAATCAAGAAATATGAGACCGAAGTAAATATGTTCTAAT 6287
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6288 TACAACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCT 6347
Qy 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6348 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTAAA 6407
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6408 TTTCAAGAAATTTTATTTCAATCAAGAAATACCTCAGGGGCTATATGTGCGAGTAGCC 6467
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6468 ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 6528 TCTTATGCTTTTGATACATAACAAAAATCTTTTAAAAATTTGGCTCTCTAAATTTAAAAAT 6587
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6588 GATAAATTCACACTATATCGGCATAGTAAAAAATACAGATATAAAGCTTTTAGAATTTCTA 6647
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6648 GAAAAAATCTTACAAAAATAAACTATATGCTTATGCTCTACAGTCITTTTAGCAAAATTT 6707
```

Db 6408 TTTCAGGAATTTATTTCAATCAAGAATTTACCTCAGGGTCTATATGTGTGACGTAGCC 6467  
Qy 141 llealaLeuGlyTyrLysGluileTyrLeuSerGlyIleAaspPheTyrClnasnGlySer 160  
Db 6468 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCA 6527  
Qy 161 SerTyrAlaPheAaspThrLysGlnGluasnLeuLeuLysLeuAlaProAaspPheLysasn 180  
Db 6528 TCTTATGCTTTTGATACATAACAAAAATCTTTTAAATTTGGCTCCTTAATTTTAAAAAT 6587  
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 6588 GATAATTCACACTATATCGGCATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 6647  
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 6648 GAAAAAATTTCAAAATAAACTATATTTGCTTATGCTTAAACAGTCTTTTAGCAATTTT 6707  
Qy 221 lleGluLeuAlaProAsnLeuAsnSerAsnPheIlelleGlnGluLysAsnAsnTyrThr 240  
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAATTAACACTACT 6767  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 6768 AAGATATACTCATACCTTCTAGTCAGGCTTATGCAAAATTTTCAAAAAATTAATTTT 6827  
Qy 261 LysLysIleLysIleLysGluasnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280  
Db 6828 AAAAAATATAAAATTAAGAAAAATATTTATTACAAGTTTGATAAAAGATCTATTAAAGATTA 6887  
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 6888 CCTAGTGATATAAGCATTTATTCAAGGAAAA 6920

RESULT 54  
AF130984  
LOCUS AF130984 11474 bp DNA linear BCT 09-FEB-2000  
DEFINITION Campylobacter jejuni lipooligosaccharide biosynthesis locus,  
partial sequence.  
ACCESSION AF130984  
VERSION AF130984.1 GI:6940827  
KEYWORDS  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
REFERENCE 1 (bases 1 to 11474)  
AUTHORS Gilbert,M., Brisson,J.R., Karwaski,M.F., Michniewicz,J.,  
Cunningham,A.M., Wu,Y., Young,N.M. and Wakarchuk,W.W.  
TITLE Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.  
Identification of the glycosyltransferase genes, enzymatic  
synthesis of model compounds, and characterization of nanomole  
amounts by 600-mhz (1)h and (13)c NMR analysis  
J. Biol. Chem. 275 (6), 3896-3906 (2000)  
JOURNAL 10660542  
PUBMED 2 (bases 1 to 11474)  
REFERENCE Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.  
AUTHORS Direct Submission  
TITLE Submitted (24-FEB-1999) Institute for Biological Sciences, National  
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A  
0R6, Canada  
FEATURES Location/Qualifiers  
source 1. 11474  
/organism="Campylobacter jejuni"  
/mol\_type="genomic DNA"  
/strain="OH4384"  
/db\_xref="taxon:197"  
<1. .360  
/note="Helicobacter pylori RfaC homolog"  
/codon\_start=1  
/transl\_table=11  
/product="heptosyltransferase I"

CDS

/protein\_id="AAF31765.1"  
/db\_xref="GI:6940828"  
/translation="KEYEPAKEVNLNSIGDIEHIELAPKENBELMAFTQWMDLIIGN  
DSGPHFLAPALNKASITIFGATPSYRNAFTQTHINKIDTCKIQNAKHIDKSDFCITR  
IREEDIFKLAKGLINEK"  
350. .1237  
/note="Helicobacter pylori WaaM homolog"  
/codon\_start=1  
/transl\_table=11  
/product="lipid A biosynthesis acyltransferase"  
/protein\_id="AAF31766.1"  
/db\_xref="GI:6940829"  
/translation="MKNSDRIYLSYLIIYKFFVTFMPDPCILFLALIVARIAFHLNKK  
HRKIINTNIQICFPQYTKERDKSLKIYENFAQYGDICLQNTTKKEKILNKYFNIN  
ENFLDALAKRPPIIFTAHYGNWEILSAYAAKYAISIVGKGLKEVMYIILQSQR  
TOFDIELDKGGRQMLKALKERALGILTDODCVENESVRLKFNKNVYNGASL  
IQRNRNALLIPVAYKEGGKFCIEFFKAKDSQNASLEELTYLQASCEEMKKRPWEY  
FFHRRFASVNEEIIYKGAK"  
1234. .2490  
/note="Neisseria meningitidis LgtF homolog"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAF31767.1"  
/db\_xref="GI:6940830"  
/translation="MNLKOISVIIIVKNAEQTLLECLNSLKDDEIILLNNESSDNTL  
KTANBFKDFANLYIVHNAFIFGALKNLALSAYAKNDWILSIDADEVLENECIELKN  
LKLQEDNIIALSRKNLYKGEWYKACGWPDYVLRIFNKNETRFENDNLVHESLVLPNSA  
KKIYKNGLUKHSYKIDISHLIDKQYSSWAKONHKSGLVKANLURAKWFFPNFYF  
LKNGLYGYKGFIIISCSALGTFPKYMKLYELQRPKPTCALIIITYNKWERLVLVD  
SVKNLAFIPNEVLADDSGDEDTARIEEYQKDFPCPLKHWOEDEGFKLSKSNKTI  
KNADSEYIIIDGDMILEKDFIKEHLEFAQRKLFLQGSRVILNKKESEIILNKDDYRI  
IFNKDFKSSKNSFLAKIFYSLSKKR"  
2786. .3955  
/note="similar to Streptococcus pneumoniae Cps14J"  
/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAF31768.1"  
/db\_xref="GI:6940831"  
/translation="MKKIGWPIYVNEKYRLRECLDSVINQTYTNLEIILVNDGSTDE  
HSLTAKETLADKRITLFDKNGGLSSARNIGIEYPSGKYLNKTKTOHIKENSILIEF  
OLDGNNPNIYKAYKSSQAPNNEKDLTNFTYPSIDYIIFLSDNYKWLNCIEECVIRM  
KQVVDVFDHDCITYEDNINKHKKTEMEIFDFKECIIITPEYANRALSVCGRDISFC  
WNGMIDFNLCQIKLKFIPFINEDIHFGIILPASANKIYVLSQLKLCIRANSIN  
HDKITKANVSEYFKDIYFTGENAKKNYLKAARKVIITALKLIEFFKQDKNENALA  
IKETFLPCVAKKALMIKFKDPLNLKEQLVLIKPFIQTQKLPYDIWKFQWKINI"  
4025. .5068  
/gene="cgtA"  
4025. .5068  
/gene="cgtA"  
/experiment="experimental evidence, no additional details  
recorded"  
/note="glycosyltransferase"  
/codon\_start=1  
/transl\_table=11  
/product="beta-1,4-N-acetylgalactosaminyltransferase"  
/protein\_id="AAF31769.1"  
/db\_xref="GI:6940832"  
/translation="MLFQSYFVKIIICLFIIPFRKIRHKIKKTFLLKNFORDKIDSYLPK  
KTLVQIKYNNEDLILKNAIIGEGHKGYENVDEKSKDPKSPLNPAFIRVKNIAITL  
KASLESILPAIORGVIGYNDCTDGSSEIILEFCQKQPSFPIKYPVEIQONPKESEN  
KLVSYNVVASPIPKDEWLIDIDHIYDAKKLYKGYIPKNKYDVVVSQRVDIHYFN  
DNFFLCKNNNGNLIKPEGDCLLINYNLKWKEVLIIRINNKKATQKQSFSSNTHSLE  
OLKYKRLIFHTLNNYHFPFLKKHRAQDIYKYKNWISIEEFKKFYQLQNHKIEPSMI  
SKETLKKIFTLF"  
complement(5054. .5959)  
/gene="cgtB"  
complement(5054. .5959)  
/gene="cgtB"  
/experiment="experimental evidence, no additional details  
recorded"  
/note="glycosyltransferase"

gene

CDS

```
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAF31770.1"
/db_xref="GI:6940833"
/translacion="MFKSIILPTYNVEQVIARAIESCINQTFKDIIEIIVVDDCGNDN
SINAKYSSKDKRIKIHNKKNLGLLRARVEGVKANSPIIMFLDPPDYELNACCE
CIIKLQDQSDVLFNAIVESNVISYKEDFNSGFYSKKEFKKIIAKKNLYWTMG
KLKKDYLEAPASLRLEKDKVIMADVDLILYPMLSQAOKIAIYMCNLYHYVPNNNS
ICNTKEVIVKNNIIQELQVLNRYLNYLNNKYCSVLIVLKIYLLYIYIKIRKTKLM
VTLAKINLTILKILPKYKKFLKQC"
```

gene

6048..6923

/gene="cst-II"

CDS

6048..6923

```
/gene="cst-II"
/experiment="experimental evidence, no additional details
recorded"
```

/note="bifunctional sialyltransferase"

/codon\_start=1

/transl\_table=11

/product="alpha-2,3/8-sialyltransferase"

/protein\_id="AAF31771.1"

/db\_xref="GI:6940834"

```
/translacion="MKKVIAGNPSLKEIDYSRLPNDVDFRCNOYFEDKYVLGKK
CKAVYNPIILFFQYVTLKHLIONQYETELIMCSNYNOAHLENENFVKFYDFPDA
HLGYDFKQKLDNAYKFKHEIYFNQRIITSGVYMCVAIALGYKEIYLSIGLDFQNGS
SYAFDTQKKNLLKLPNFRNDSHYIGHSKNNTDIALEFLEKTKIKLYCLCPNSLLA
NFIELAPNLSNFTIQKNNTYTKDILIPSEAYGKFSKNINFKKIKIKENIYYKLKD
LLRLPSDIKHVFYKKG"
```

CDS

6924..7964

/note="Neisseria meningitidis SiaC homolog"

/codon\_start=1

/transl\_table=11

/product="sialic acid synthase"

/protein\_id="AAF31772.1"

/db\_xref="GI:6940835"

```
/translacion="MKKIKIONIIISBEKAPLVVPEIYNHNGSLELAKIMVDAFST
GAKIKRQTHIVEDMSKAQKVIIPGNAKISIVIEIMOKCALDYKDELAKYEYKELGL
VYLSVTPFSRAGANLEDMVSAFKIGSGECNNYPLIKHIAFKPMIVSHDTMNSIESI
KPTVKILLNDEIPVLMTNLYPTPHNLVRLNAMLKELKFEFCMVLGSLSDHTGNLAC
LGAVALGCVLERHFTDSMHSRGPDIVCSMDTQALKEIILIQSEQMAIMRGNESKKA
KQEVTTIDFAFASVVSIIKIKGBVLSMDNIWVRPGLGGISAAEPENILGKKALRDI
ENDTQLSYEDFA"
```

CDS

8021..9079

```
/note="involved in sialic acid biosynthesis; Neisseria
meningitidis SiaA homolog"
```

/codon\_start=1

/transl\_table=11

/product="sialA-like protein"

/protein\_id="AAF31773.1"

/db\_xref="GI:6940836"

```
/translacion="MYRVQNSSEFELYIFATGMHLSNFGYTVKELYKNGFKNITYEPI
NDKYFSTDLALHAIIDGFRYVYNELKPDILIVHGDRIEPLAAIIVGALANNILVAHIE
GSGTSGTIDSLRAHISKLAHILVNDDEFARLMQIGEDKSFIFIGSPDLRLNDN
KISLNEAKKYDINVENYALLMHPVTEITSIKNQADNLVKALIQSNKVIYIYPNN
DLGFLILQSYEEUNPKFLPSLRFEYFTLLKNADFIIGNSCILKEALYIKTA
GILVGSQRGNLGNENTKVNANSDEILKAINTIHKQDLFSAKLEILDSSKLPFEYL
QSGBFFKLNITQKVFEDIK"
```

CDS

9076..9741

/note="Haemophilus ducreyi NeuA homolog"

/codon\_start=1

/transl\_table=11

/product="CMP-sialic acid synthetase"

/protein\_id="AAF31774.1"

/db\_xref="GI:6940837"

```
/translacion="MSLAIIPARGSGKIKNNKLVLLNNKPLIYVYTIKAALNTKSISK
VWVSSDSEILNYSQNDILKRPISLAQNTTSDKVLHLAKFYKDIEDVFLQPT
SPLRNTHIDEAFNLYKSNALISVSCENKILKAFVNCYEGDLAGIENDEYFPMP
RQKLPTKYSNGAIYILKIEFLNPNPFLQSKTKHFLMDESSSLDIDCLLEDLKKAEQI
WKK"
```

CDS

9729..10562

```
/note="similar to Rhodobacter capsulatus hypothetical
protein encoded by GenBank Accession Number AF010496"
```

```
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAF31775.1"
/db_xref="GI:6940838"
/translacion="MEKITLCKNKNILNLKQYNIYTKYIENPRFSLTKTKDFITF
PNNQPHDRISTASFTYETNHSFINDAQONHINKTFIVNHNPSSSIHLIIQDDVWI
FKDVLAKQGITLGTGCVIGORAVTKDVPVAIVAGIPAKIIKYRDEKTIERLLKIQ
WKVHPADFYDIDLNKINQYDLLEEKIIKKSISYYPNPKLYFRDILLELKSXKIFNL
F"
complement(10554..11366)
/note="similar to Haemophilus influenzae hypothetical
protein HI0868 encoded by GenBank Accession Number U32768"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAF31776.1"
```

CDS

11474

```
Alignment Scores:
Pred. No.: 5,55e-184 Length: 11474
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 4
Query Match: 97.6% Indels: 0
DB: 15 Gaps: 0
```

US-10-734-719-9 (1-291) x AFI30984 (1-11474)

Qy	1	MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg	20
Db	6048	ATGAAAAAGTTATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTCATTATTCAAGA	6107
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	6108	CTACCAATGATTTTGTATGATGATATCAATTTTATTATTGAAGATAAATACTAT	6167
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	6168	CTTGGTAAAAAATGCAAGCGAGTATTTTACAACTCTATTCTTTTTTTTGAACAACTACTAC	6227
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	6228	ACTTTAAACATTTAATCCAAAAATCAAGAAATATAGACCGAATAATATATGTTCTAAT	6287
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	6288	TACAACCAAGCTCATCTAGAAAAATGAAAAATTTTGTAAAAACTTTTTCACGATTATTTTCCT	6347
Qy	101	AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	6348	GATGCTCATTTGGGATATGATTTTTTCAACAACTATAAGATTTTAATGCTTATTTTAA	6407
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	6408	TTTACGAAATTTATTTCAATCAAGAATTACTCAGGGTCTATATGTCAGTAGCC	6467
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	6468	ATAGCCCTAGGATACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCA	6527
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180
Db	6528	TCATTATGCTTTTGATCTACAAACAAAAATCTTTTTAAAAATTCGCTCCTAAATTTAAAAAT	6587
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	6588	GATTAATTCACACTATATCGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTA	6647
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	6648	GAAAAAATCTACAAAAATAAACTATATTGCTTATGCTTACACAGTCTTTTAGCAATTTT	6707

```

QY 221 ileGluLeuAlaProAsnLeuAsnSerAsnPheIleileGlnGluLysAsnTyrThr 240
Db 6708 ATAGAACTAGCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAAATAACTACACT 6767
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerIlyshsnIleAsnPhe 260
Db 6768 AAGAGATATACATACCTTCTAGTGAGGCTTATGGAAAAATTTTCAAAAAATATTAATTTT 6827
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6828 AAAAAAATAAATAATTAAGAAAAATATTTATTCAAGTTGATTAAGATCTATTAAAGATTA 6887
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6888 CCTAGTATATAAGCAATTTTCAAGGAAAA 6920

RESULT 55
AY297047
LOCUS AY297047 12576 bp DNA linear BCT 28-JUL-2003
DEFINITION Campylobacter jejuni HB93-13 wlaII LOS biosynthesis gene cluster,
partial sequence.
ACCESSION AY297047
VERSION AY297047.1 GI:33151090
KEYWORDS
SOURCE
ORGANISM
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 12576)
Nawagamuwa, V., Coloe, P.J. and Fry, B.N.
The wlaII LOS biosynthesis locus of C. jejuni HB93-13 (HS:19)
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 12576)
Nawagamuwa, V., Coloe, P.J. and Fry, B.N.
Direct Submission
JOURNAL Submitted (12-MAY-2003) Biotechnology & Environmental Biology,
Royal Melbourne Institute of Technology University, Plenty Road,
Bundoora, VIC 3083, Australia
FEATURES
source
1..12576
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="HB93-13"
/serotype="HS:19"
/db_xref="taxon:197"
<1..>12576
misc_feature
/note="contains wlaII LOS biosynthesis gene cluster"

ORIGIN
Alignment Scores:
Pred. No.: 6.06e-184 Length: 12576
Score: 1525.00 Matches: 283
Percent Similarity: 98.6% Conservative: 4
Best Local Similarity: 97.3% Mismatches: 0
Query Match: 97.6% Indels: 0
Gaps: 0

US-10-734-719-9 (1-291) x AY297047 (1-12576)

QY 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6063 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAAGA 6122
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6123 CTACCAATGATTGATGATTATTAGATGTAATCAATTTTATTGAGATAAATACTAT 6182
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6183 CTTGTGTAATAAATGCAAGGCAGTATTTTACAATCCTATTCTTTTGTGAAACAATACTAC 6242
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80

```

```

Db 6243 ACTTTAAACATTTAATCCAAATCAAGATATAGAGACCGCAACTAATTTATGTGTTCTAAT 6302
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6303 TACACCAAGCTCATCTAGAAAATGAAAATTTGTAAAAAATCTTTTACGATTATTTTCTT 6362
QY 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6363 GATGCTCATTTGGATATGATTTTTCACAAACACTTAAAGATTTTAATGCTTATTTTAAA 6422
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6423 TTTTACGAAATTTTATTTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGTGCGATGCC 6482
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6483 ATAGCCCTAGGATACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGCA 6542
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6543 TCTTATGCTTTTGATACTAAACAAAAATCTTTTAAATTTGGCTCCTAATTTTAAAAAT 6602
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6603 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTA 6662
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6663 GAAAAAACTTACAAAAATAAACTATATTTGCTTATGCTTCTTAAACAGTCTTTTAGCAAAATTT 6722
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleileGlnGluLysAsnAsnTyrThr 240
Db 6723 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATAACTACTACT 6782
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6783 AAGATATATCTACATCTCTCTAGTGGCTTATGGAATAATTTTCAAAAAATATTAATTTT 6842
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6843 AAAAAAATAAATAATTAAGAAAAATTTTATTAACAAGTTGATAAAGATCTATTAAAGATTA 6902
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6903 CCTAGTATATAAGCAATTTTCAAGGAAAA 6935

RESULT 56
AY681242
LOCUS AY681242 1303 bp DNA linear BCT 06-MAY-2005
DEFINITION Campylobacter jejuni isolate Contig_16b Cj81-010 and Cj81-011
genes, partial cds.
ACCESSION AY681242
VERSION AY681242.1 GI:57790356
KEYWORDS
SOURCE
ORGANISM
Campylobacter jejuni subsp. jejuni 81-176
Campylobacter jejuni subsp. jejuni 81-176
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 1303)
Poly, F., Threadgill, D. and Stintzi, A.
Genomic diversity in Campylobacter jejuni: identification of C.
jejuni 81-176-specific genes
JOURNAL J. Clin. Microbiol. 43 (5), 2330-2338 (2005)
PUBMED 15872262
REFERENCE
2 (bases 1 to 1303)
Poly, F. and Stintzi, A.
Direct Submission
JOURNAL Submitted (08-JUL-2004) Veterinary Pathobiology, Oklahoma State
University College of Veterinary Medicine, 250 McElroy Hall,
Stillwater, OK 74078, USA
FEATURES
Location/Qualifiers
source
1..1303
/organism="Campylobacter jejuni subsp. jejuni 81-176"

```

```
/mol_type="genomic DNA"
/strain="81-176"
/isolate="Contig_16b"
/db_xref="taxon:354242"
complement(<1..361)
/notes="similar to beta-1,3-galactosyltransferase in
Campylobacter jejuni OH4384"
/codon_start=1
/transl_table=11
/product="Cj81-010"
/protein_id="AAW56103.1"
/db_xref="GI:57790357"
/translation="MKISILPTYNVEQVIARAIESCINQTKDIEIIVVDDGNDN
SINIAEYKSKDKRIKIIHNEKNLGLURAYEGVKVANSPIYIMFLDPPDYLELNACEE
KILDEQDEVDLVFFNA"
450..>1303
/notes="similar to alpha-2,3-sialyltransferase in
Campylobacter jejuni OH4384"
/codon_start=1
/transl_table=11
/product="Cj81-011"
/protein_id="AAW56104.1"
/db_xref="GI:57790358"
/translation="MKVIIAGNGPSLKEIDYSLPNDPVERCNOFYFEDKYVLGKK
CKAVFTPNFFEQYITLKHILQNOEYETELIMCSYNVNOAHLENENFVKTFDYFPPA
HLGYDFPKQJKEFNAYPKFHEIYFNQRTISGVNCAVAIALGYKEIYLSGIDFQNGS
SYAFDTKQENLLKLPDKNDRSHYIGHSKNTDIKALEFLEKTYKIKLYCLCPNSLLA
NFIELAPNLNSNFIQEKNNYTKDILIPSSSEAYGKFSKNINFKIKIKENYVYKLKD
LLRLPSDI"
ORIGIN
Alignment Scores:
Pred. No.: 2,21e-184 Length: 1303
Score: 1521.00 Matches: 284
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.3% Indels: 0
DB: 15 Gaps: 0

US-10-734-719-9 (1-291) x AY681242 (1-1303)

Qy 1 MetLysLysValIleIleAlaGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 450 ATGAAAAAAGCTTATATGCTGGAATGGCAAGTTTAAAGAAATGGATTATCAAGG 509
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 510 CTACCAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 569
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 570 CTTGGTAAAAATGCAAGCAGTGTGTTTACACCCCTAATTTCTTTTGGAGATACTAC 629
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 630 ACTTTAAACATTTAAATCCAAATCAAGATATGAGACCGAATTAATGATGTTCTAAT 689
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 690 TACACCAAGCTCATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTATTTCTCT 749
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 750 GATGCTCATTTGGATATGATTTTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAA 809
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 810 TTTTCAAGAAATTTATTTCAATCAAGAAATTAACCTCAGGGGTCTATATGTCGCGATGCC 869
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 870 ATAGCCCTAGATACAAAGAAATTTATCTTTCCGGAAATTTGATTTTATCAAAATGGGTCA 929
```

```
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 930 TCTTATGCTTTTATACCAACCAAGAAATCTTTTAAACTAGCCCTGATTTTAAAAAT 989
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 990 GATCGCTCACACTATATCGGACATAGTAAATAATACAGATATATAAGCTTTAGAAATTTCTA 1049
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 1050 GAAAAAATCTTACAAAAATAAACTATATTCCTTATGCTCTAATAGTCTTTTAGCAAAATTTT 1109
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 1110 ATAGAACTAGCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAATAACTACTACT 1169
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 1170 AAAGATATACTCATACCTTCTAGTCAGGCTTATGAAAAATTTTCAAAAAATATAATTTT 1229
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeuArgLeu 280
Db 1230 AAAAAAATAAAAAATTAAGAAATGTTTATTACAAAGTTTGATAAAAGATCTATTAGATTA 1289
Qy 281 ProSerAspIle 284
Db 1290 CCTAGTGATATA 1301
RESULT 57
AX934429
LOCUS AX934429 876 bp DNA linear PAT 05-JAN-2004
DEFINITION Sequence 6 from Patent WO02074942.
ACCESSION AX934429
VERSION AX934429.1 GI:40641676
KEYWORDS
SOURCE
ORGANISM
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: WO 02074942-A 6 26-SEP-2002;
National Research Council of Canada (CA)
FEATURES
source
1..876
/organism="Campylobacter jejuni"
/mol_type="unassigned DNA"
/db_xref="taxon:197"
1..876
/note="unnamed protein product; Campylobacter
alpha-2,3/alpha 2,8-sialyltransferase II (cstII) from C.
jejuni serotype O:41"
/codon_start=1
/transl_table=11
/protein_id="CAF04168.1"
/db_xref="GI:40641677"
/translation="MKVIIAGNGPSLKEIDYSLPNDPVERCNOFYFEDKYVLGKK
CKAVFNPFLFEQYITLKHILQNOEYETELIMCSNFNOAHLENONFVKTFDYFPPA
HLGYDFPKQJKEFNAYPKFHEIYFNQRTISGVNCAVAIALGYKEIYLSGIDFQNGS
SYAFDTKQENLLKLPAPNPNNDSHYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLA
NFIELAPNLNSNFIQEKNNYTKDILIPSSSEAYGKFTKNINFKIKIKENYIYKLIKD
LLRLPSDIKHVFKG"
```

## ORIGIN

## Alignment Scores:

```
Pred. No.: 1,18e-183 Length: 876
Score: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservative: 8
Best Local Similarity: 95.9% Mismatches: 4
Query Match: 96.9% Indels: 0
DB: 2 Gaps: 0
```

US-10-734-719-9 (1-291) x AX934429 (1-876)

```
QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
DB 1 ATGAAAAAGTTATTATTGCTGGAATGGCAAGCTTTAAAGAAATTTGATTATTTCAAGA 60
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
DB 61 CTACCAATGATTTTGTATGATTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
DB 121 CTGTGTAATAAATGCAAGCAGTATTACCAATCTAGTCTTTTTCGAACAATACTAC 180
QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
DB 181 ACTTTAAACAATTTAATCAAAATCAAGAATATGAGCCGAATCAATCATGTGTTCTAAT 240
QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
DB 241 TTTTAAACCAAGCTCATCTAGAAATCAAAATTTTGTAAACACTTTTACGATTATTTTCT 300
QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
DB 301 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAGAATTTCAATGCTTATTTTAA 360
QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
DB 361 TTTTCAGAAATTTATTCAATCAAGAATTTACCTCAGGGGTCTATATGTGCAGTAGCC 420
QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGCA 480
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
DB 481 TCTTATGCTTTTGTATACCAACAAACAAATCTTTTAAATTTGGCTCTCTAATTTAA 540
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
DB 541 GATAATTCACACTATATCGACATAGTAAATAATACAGATATATAAGCTTTTAGAATTTCTA 600
QY 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
DB 601 GAAAAAACTTACGAATAAAGCTATATTTGTTATGTCCTTAACAGTCTTTTAGCAAAATTT 660
QY 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnSerTyrThr 240
DB 661 ATAGAACTAGCGCCAAATTTAAATTTCAAAATTTTATCATCAAGAAAAAATAACTAT 720
QY 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
DB 721 AAAGATATACCTATACCTTCTAGTGAGGCTTATGGAATTTTACAAAAATATTAAATTT 780
QY 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
DB 781 AAAAAATATAAAATTTAAAGAAATTTATTATCAAGTTGATATAAGATCTATAAGATTA 840
QY 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
DB 841 CCTAGTGATATAAGCATTTATTTCAAGGAAAA 873
```

RESULT 58

LOCUS

BD249793

DEFINITION

Campylobacter glycosyltransferases for biosynthesis of

gangliosides and ganglioside mimics.

ACCESSION

BD249793

VERSION

BD249793.1 GI:33059563

KEYWORDS

JP 2002535992-A/4.

SOURCE

Campylobacter jejuni

ORGANISM

Campylobacter jejuni

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.

1 (bases 1 to 876)

Gilbert, M. and Wakarchuk, W.W.

Campylobacter glycosyltransferases for biosynthesis of

gangliosides and ganglioside mimics

Patent: JP 2002535992-A 4 29-OCT-2002;

NATIONAL RESEARCH COUNCIL OF CANADA

OS Campylobacter jejuni

PN JP 2002535992-A/4

PD 29-OCT-2002

PF 01-FEB-2000 JP 2000597438

PR 01-FEB-1999 US 60/118213, 31-JAN-2000 US 06/495406 P1

MICHEL GILBERT, WARREN W WAKARCHUK

PC C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00 CC

Campylobacter glycosyltransferases for biosynthesis of

gangliosides and

CC ganglioside mimics

FH Key

FT source

1. .876

Location/Qualifiers

/organism='Campylobacter jejuni'

/mol\_type='genomic DNA'

/db\_xref='taxon:197'

ALIGNMENT SCORES

Pred. No.: 1.18e-183

Score: 1514.00

Length: 876

Matches: 279

Percent Similarity: 98.6%

Conservative: 8

Best Local Similarity: 95.9%

Mismatches: 4

Query Match: 96.9%

Indels: 0

Gaps: 0

DB: 2

US-10-734-719-9 (1-291) x BD249793 (1-876)

QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20

DB 1 ATGAAAAAGTTATTATTGCTGGAATGGCAAGCTTTAAAGAAATTTGATTATTTCAAGA 60

QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

DB 61 CTACCAATGATTTTGTATGATTTTAGATGCAATCAATTTTATTTGAAGATAAATACTAT 120

QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

DB 121 CTGTGTAATAAATGCAAGCAGTATTACCAATCTAGTCTTTTTCGAACAATACTAC 180

QY 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80

DB 181 ACTTTAAACAATTTAATCAAAATCAAGAATATGAGCCGAATCAATCATGTGTTCTAAT 240

QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100

DB 241 TTTTAAACCAAGCTCATCTAGAAATCAAAATTTTGTAAACACTTTTACGATTATTTTCT 300

QY 101 AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

DB 301 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAGAATTTCAATGCTTATTTTAA 360

QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

DB 361 TTTTCAGAAATTTATTCAATCAAGAATTTACCTCAGGGGTCTATATGTGCAGTAGGCC 420

QY 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

DB 421 ATAGCCCTAGGATACAAAGAAATTTATCTTTCGGAAATTTGATTTTATCAAAATGCA 480

QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180

DB 481 TCTTATGCTTTTGTATACCAACAAACAAATCTTTTAAATTTGGCTCTCTAATTTAA 540

```

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTACACACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTACGAAATAAAGCTATATGTTATGTCCTAACAGCTCTTTAGCAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTATATACAAAGAAAAATAAATACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTATACACTCTTAGTACGAGCTTATGGAATAATTTACAAAAATATTAATTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAATAAATAAAGAAATATTTATACAAAGTTGATAAAGATCTATTAAAGATTAA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATATTATTTCAAAGGAAAA 873

RESULT 59
LOCUS AR271702 876 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 6 from patent US 6503744.
ACCESSION AR271702
VERSION AR271702.1 GI:29703247
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6503744-A 6 07-JAN-2003;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.18e-183 Length: 876
Pred. No.: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservativity: 8
Best Local Similarity: 95.9% Mismatches: 4
Query Match: 96.9% Indels: 0
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR271702 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAGTTATTTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATCAAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGATTTTAGTGCATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTTGGTAAAAAATGCAAGCAGTATTTTACAAATCCTAGTCTTTTTTTTGAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACITTAACAAATTTAATCCAAATCAAGATATGAGACCGAAGCTAATCATGTGTTCTAAT 240

```

```

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAAACCAGCTCATCTAGAAAATCAAATTTTGTAAAAACTTTTACGATTATTTCTT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGCATATGATTTTTCACAACTTAAAGAAATTCAAATGCTTATTTTAAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAAGAAATTTATTTCAATCAAAGAAATTAACCTCAGGGGTCTATATGTGCACAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGATCA 480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 481 TCTTATGCTTTTGATACCAACAAATAATCTTTTAAAAATTTGGCTCCTAAATTTAAAAAT 540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 541 GATAATTACACACTATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTA 600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 601 GAAAAAATTACGAAATAAAGCTATATTTATGTTATGCTCTAACAGTCTTTTACCAAAATTTT 660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 661 ATAGAACTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAAGAAAAATAAATACTACT 720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 721 AAAGATATATCTATACACTCTTAGTACGAGCTTATGGAATAATTTACAAAAATATTAATTTT 780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 781 AAAAAATAAATAAATAAAGAAATTTATTTACAAAGTTGATAAAGATCTATTAAAGATTAA 840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 841 CCTAGTATATAAGCATATTATTTCAAAGGAAAA 873

RESULT 60
LOCUS AR481784 876 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 6 from patent US 6699705.
ACCESSION AR481784
VERSION AR481784.1 GI:47243419
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 6 02-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;

FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.18e-183 Length: 876
Pred. No.: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservativity: 8
Best Local Similarity: 95.9% Mismatches: 4

```

Query Match: 96.9% Indels: 0  
DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR481784 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20  
Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTTCAAGA 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40  
Db 61 CTACCAATGATTTTGGTATGATTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60  
Db 121 CTGTGTAATAAATGCAAGCAGTATTTTACAACTCTAGTCTTTTGTGAACAATACTAT 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80  
Db 181 ACTTTAAACATTTTAACTCAAAATCAAGAATATGAGACCGAACTAATCATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
Db 241 TTTAACCAAGCTCATCTAGAAATTTGTAAAACTTTTACGATTTATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
Db 301 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGAAATTTCAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCACAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160  
Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGATCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180  
Db 481 TCTTATGCTTTTGATACCAACAAAAAATCTTTTAAAAATTTGGCTCTCTAATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200  
Db 541 GATAATTCACACTATATCGGACATAGTAAATAATACAGATATAAAAGCTTTTGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLysLeuTyrCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 601 GAAAAAATTTAGCAATAAAGCTATATTTTATGTCCTTAACAGTCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTATACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260

## RESULT 61

AR527383

LOCUS

DEFINITION Sequence 6 from patent US 6723545.

ACCESSION AR527383

VERSION AR527383.1 GI:53914404

KEYWORDS

SOURCE

linear

PAT 08-OCT-2004

876 bp DNA

Unknown.

## ORGANISM

Unknown.

Unclassified.

1 (bases 1 to 876)

AUTHORS Gilbert, M. and Wakarchuk, W. W.

Polypeptides having .beta.-1,4-GalNAC transferase activity

Patent: US 6723545-A 6 20-APR-2004;

National Research Council of Canada; Ottawa;

CAX;

Location/Qualifiers

1..876

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:

Pred. No.: 1.18e-183 Length: 876

Score: 1514.00 Matches: 279

Percent Similarity: 98.6% Conservative: 8

Best Local Similarity: 95.9% Mismatches: 4

Query Match: 96.9% Indels: 0

DB: 2 Gaps: 0

US-10-734-719-9 (1-291) x AR527383 (1-876)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArg 20

Db 1 ATGAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAAAGAAATTTGATTATTTCAAGA 60

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

Db 61 CTACCAATGATTTTGGTATGATTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT 120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60

Db 121 CTGTGTAATAAATGCAAGCAGTATTTTACAACTCTAGTCTTTTGTGAACAATACTAT 180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuMetCysSerAsn 80

Db 181 ACTTTAAACATTTTAACTCAAAATCAAGAATATGAGACCGAACTAATCATGTGTTCTAAT 240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100

Db 241 TTTAACCAAGCTCATCTAGAAATTTGTAAAACTTTTACGATTTATTTCTCT 300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120

Db 301 GATGCTCATTTGGGATATGATTTTTCACCAACTTAAAGAAATTTCAATGCTTTATTTAAA 360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140

Db 361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTCACAGTAGCC 420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

Db 421 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGATCA 480

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180

Db 481 TCTTATGCTTTTGATACCAACAAAAAATCTTTTAAAAATTTGGCTCTCTAATTTTAAAAAT 540

Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200

Db 541 GATAATTCACACTATATCGGACATAGTAAATAATACAGATATAAAAGCTTTTGAATTTCTA 600

Qy 201 GluLysThrTyrLysIleLysLysLeuTyrCysProAsnSerLeuLeuAlaAsnPhe 220

Db 601 GAAAAAATTTAGCAATAAAGCTATATTTTATGTCCTTAACAGTCTTTTAGCAAAATTTT 660

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240

Db 661 ATAGAACTAGCGCCAAATTTAAATTTCAATTTTATCATACAGAAAAAATACTATACT 720

Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260



```
|||||
721 AAAGATATACCTCATCTCTAGTGGCTTATGGAAATTTACAAAATAATAATTT 780
Db

261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysAspLeuLeuArgLeu 280
Qy
|||||
781 AAAAAATAAAAAATTAAGAAAAATATTTATTACAAAGTTGATAAAAGATCTATTAAGATTA 840
Db

281 ProSerAspLeuLysHisTyrPheLysGlyLys 291
Qy
|||||
841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873
Db

RESULT 62
AR609662
LOCUS AR609662 876 bp DNA linear PAT 15-DEC-2004
DEFINITION Sequence 6 from patent US 6825019.
ACCESSION AR609662
VERSION AR609662.1 GI:56664962
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,3-galactosyl transferase activity
JOURNAL Patent: US 6825019-A 6 30-NOV-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1..18e-183 Length: 876
Score: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservative: 8
Best Local Similarity: 95.9% Mismatches: 4
Query Match: 96.9% Indels: 0
Gaps: 2
US-10-734-719-9 (1-291) x AR609662 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 61 CTACCAATGATTTTGATGATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 121 CTGGTAAAAATGCAAAAGCAGTATTTTACAACTCTAGTCTTTTGTGAAACAATACTAC 180
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 181 ACTTTAAACAATTTAATCCAAAATCAAGAAATATGAGACCGAACTAATCATGTGTCTAAT 240
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 241 TTTAACCAAGCTCATCTAGAAAATCAAAAATTTTGTAAGAACTTTTACGATTATTTTCT 300
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 301 GATGCTCATTTGGATATGATTTTTCACAACTTAAAGAAATTCATGCTTATTTTAA 360
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 361 TTTCAACGAAATTTATTTCAATCAAAAGAAATACCTCAGGGGTCTATATGTCACAGTAGCC 420
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
|||||
421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGATCA 480
Db

161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Qy
|||||
481 TCTTATGCTTTTGATACCAAAACAAAATCTTTTAAAAATGGCTCCTTAATTTTAAAAAT 540
Db

181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Qy
|||||
541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db

201 GlulysThrTyrLysLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Qy
|||||
601 GAAAAAATCTACGAAATAAAGCTATATTGTTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660
Db

221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Qy
|||||
661 ATAGAACCTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTATACT 720
Db

241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Qy
|||||
721 AAAGATATACCTCATCTCTAGTGGCTTATGGAAATTTACAAAATAATAATTTT 780
Db

261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu 280
Qy
|||||
781 AAAAAATAAAAAATTAAGAAAAATATTTATTACAAAGTTGATAAAAGATCTATTAAGATTA 840
Db

281 ProSerAspLeuLysHisTyrPheLysGlyLys 291
Qy
|||||
841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873
Db

RESULT 63
AR689940
LOCUS AR689940 876 bp DNA linear PAT 12-SEP-2005
DEFINITION Sequence 6 from patent US 6905867.
ACCESSION AR689940
VERSION AR689940.1 GI:74471948
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Nucleic acids encoding polypeptides with .beta.1,3-galactosyl transferase activity
JOURNAL Patent: US 6905867-A 6 14-JUN-2005;
National Research Council of Canada; Ottawa;
CAX;
FEATURES
source Location/Qualifiers
1..876
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1..18e-183 Length: 876
Score: 1514.00 Matches: 279
Percent Similarity: 98.6% Conservative: 8
Best Local Similarity: 95.9% Mismatches: 4
Query Match: 96.9% Indels: 0
Gaps: 2
US-10-734-719-9 (1-291) x AR689940 (1-876)
Qy 1 MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 1 ATGAAAAAAGTTATTATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTATTCAGA 60
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 61 CTACCAATGATTTTGATGATTTAGATGCAATCAATTTTATTTTGAAGATAAATACTAT 120
Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
|||||
421 ATAGCCCTAGGATACAAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGATCA 480
Db

161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Qy
|||||
481 TCTTATGCTTTTGATACCAAAACAAAATCTTTTAAAAATGGCTCCTTAATTTTAAAAAT 540
Db

181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Qy
|||||
541 GATAATTCACACTATATCGGACATAGTAAAAATACAGATATAAAAGCTTTAGAAATTTCTA 600
Db

201 GlulysThrTyrLysLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Qy
|||||
601 GAAAAAATCTACGAAATAAAGCTATATTGTTTATGTCCTAACAGTCTTTTAGCAAAATTTT 660
Db

221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnTyrThr 240
Qy
|||||
661 ATAGAACCTAGCGCAAAATTTAAATTTCAAAATTTTATCATACAGAAAAAATAACTATACT 720
Db

241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Qy
|||||
721 AAAGATATACCTCATCTCTAGTGGCTTATGGAAATTTACAAAATAATAATTTT 780
Db

261 LysLysLysLysLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu 280
Qy
|||||
781 AAAAAATAAAAAATTAAGAAAAATATTTATTACAAAGTTGATAAAAGATCTATTAAGATTA 840
Db

281 ProSerAspLeuLysHisTyrPheLysGlyLys 291
Qy
|||||
841 CCTAGTGATATAAGCAATTTTCAAAAGGAAAA 873
Db
```



cds; lipid A biosynthesis acyltransferase (htrB), putative two-domain glycosyltransferase, putative glycosyltransferase, putative beta-1,4-N-acetylgalactosaminyltransferase (cgtA), putative beta-1,3-galactosyltransferase (cgtB), alpha-2,3-/alpha-2,8-sialyltransferase (cstII), putative sialic acid synthase (neuB1), putative N-acetylglucosamine-6-phosphate (neuC1), CMP-Neu5Ac synthetase (neuA1), putative acetyltransferase, and putative glycosyltransferase (waaV) genes, complete cds; and heptosyltransferase II (waaF) gene, partial cds.

ACCESSION AY044868

VERSION 1

KEYWORDS

SOURCE

ORGANISM

Campylobacter jejuni

Campylobacter jejuni

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

Campylobacteraceae; Campylobacter.

REFERENCE 1 (bases 1 to 11455)

AUTHORS Gilbert M., Karwaski M.F., Bernatchez S., Young N.M., Taboada E.,

Michniewicz J., Cunningham A.M. and Wakarchuk W.W.

TITLE The Genetic Bases for the Variation in the Lipo-oligosaccharide of

the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF

SIALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE

J. Biol. Chem. 277 (1), 327-337 (2002)

PUBMED 11689567

REFERENCE 2 (bases 1 to 11455)

AUTHORS Gilbert M., Michniewicz J., Karwaski M.-F., Cunningham A. and

Wakarchuk W.W.

TITLE Direct Submission

Submitted (10-JUL-2001) Institute for Biological Sciences, National

Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A

0R6, Canada

FEATURES

source

1. 11455

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strain="ATCC 43460"

/db\_xref="ATCC:43460"

/db\_xref="taxon:197"

/notes="Penner-type strain O:41"

<1.348

/gene="waaC"

<1.348

/gene="waaC"

/codon\_start=1

/transl\_table=11

/product="heptosyltransferase I"

/protein\_id="AAK95995.1"

/db\_xref="GI:15430477"

/translation="FAKEVLNLGSDITHIELAPKPSLEELMAFTKMDLLIGNDSGP

THLAFALNKASITIFGATPSYRNAFQTHINKIIDAGKKIQNTXHKIDKSDFCITRIEE

DIFPLAKGLLNEX"

338..1225

/gene="htrB"

338..1225

/gene="htrB"

/codon\_start=1

/transl\_table=11

/product="lipid A biosynthesis acyltransferase"

/protein\_id="AAK95996.1"

/db\_xref="GI:15430478"

/translation="MKNSDRILYSLYILKFFVTFMPECEILHFLALIVARITFVLNKK

HRKIIDTNLQICFPQTKERDKSLKIENFAQFDCIQNTTKERILNKVNFN

ENLIDALAKRPITFTTHYGNWEILSLAYAAKYGAIISVKGKLSKEIMYELISQSR

TQFDIELDKGGIROMLSALKWEIRALGILTDQCVENESVRLKFNKVEYQMGASL

IAORSNALIIPVAYKEDGKFCIEFFKAKDSQNASLEELTYLQAQSCCEIMKKRPWEY

FFPHRRASVDKEIYKAK"

1222..2769

/note="orf3"

/codon\_start=1

/transl\_table=11

/product="putative two-domain glycosyltransferase"

/protein\_id="AAK95997.1"

/db\_xref="GI:15430479"

/translation="MNLKQISVIIIVKNAEQTLLECLNSLKDFDIBIILLNNESSDNTL

KIANEFKKNPANLHIYVSATFGKALKNLAYSANWILSDIDADLENEBICELKN

LEIQEDNIIALSRLNLYKGWIKACGMPDVLRIENKNTFRFNDLSLVHESILVPGNV

KKHLKNGLSHYAFRDSLHLINKQYSSIAKQNIHKSGVLKANLAFWTFERNYF

LKNGFLYGGFTISVCNAGTGFYKMKLYELQKQPKTCALITTYQKERLKLVD

SVKNLFLPNEVLIAADDSKEDTARLIEYQKDFPCLSHIHQWEDGEFPAKSRNKAI

AFSKSEYIIILIDGMILEKDFIKNHFFAKRVILQGSRIILNKESEIEILKNDYSL

AFSKDFKSVKNSFLAKIYKIKIDQKFNTKEIIKGSKTCNMFPTDFDKLDGFN

ENFIMGREDSEFVARFLFNKGIIFRLKFAIAYHIYHEENSKMLESNHQIYLDITIK

NKLSWR"

2773..3966

/note="orf4"

/codon\_start=1

/transl\_table=11

/product="putative glycosyltransferase"

/protein\_id="AAK95998.1"

/db\_xref="GI:15430480"

/translation="MKKIGVPIPIVNEKYLRECLDSVINQTYTNLEIMLVNDGSTDE

NSLNIARKYTLKDRITLFDKNGGLSSARNVGEYFSGYKLNKTKHQIKENSITIEF

QLDGNNPIYKAYKSSQAFNNEKDLNTFTYPSIDYIIFLSDNDYWLNCIEBCEVIRM

KNDVLWFDHDCYEDNINKHKTRMEIIPNKEECITPKKEYANRLSVGSRDISFG

WGMIDFNFLKQIKLKFNFILNEDIHFGIILFASANKIYVLSQLYLCRLRANSISN

HDKKITRANVSEYFDIYETEGENAKENYDKAASRVITALKLIEFFKDKQKSENALA

IKETFLPCYAKKALMIKKFKDPLNLKEQLVLIRPFITQIKUPYDIWKFWQKIKIFNN

KNIKN"

4015..5058

/gene="cgtA"

4015..5058

/gene="cgtA"

/codon\_start=1

/transl\_table=11

/product="putative

beta-1,4-N-acetyl-galactosaminyltransferase"

/protein\_id="AAK95999.1"

/db\_xref="GI:15430481"

/translation="MLFQSYEVKICLFIPFRKIRHKIKTKTFLKNIQRDKIDSYLPLK

KTAAQINKYNNEDLIKLNKAIIGBEHKGYNVDEKSKDPKPLNFAWTRVKNEAITL

KASLESILPAIQGVIGVYDGSSEIILFCKQYPSFPIKYVEIQTQPKSEEN

KLISYNNVASFIPKDEWLKIDVDHIYDAKKLYKSYFIPKNKYDVVSRYVDIHYFN

DNFFLCKDNNNGNIDKEPGDCLLINNNLKWKEVLIDRINNWKKATKQSFSSNISLE

QLKYKRIILPHTELNNYHFFFLKXKRAQDIYKYNWISIEEPFKFYQLQNIHNKIEPSMI

SKETKKIFLTLF"

complement (5044..5949)

/gene="cgtB"

complement (5044..5949)

/gene="cgtB"

/codon\_start=1

/transl\_table=11

/product="putative beta-1,3-galactosyltransferase"

/protein\_id="AAK96000.1"

/db\_xref="GI:15430482"

/translation="MFKISIIILPTVNEQYIARAIESCINQTFKDIIEIIVDDCGNDN

SINTAKESKKDKRIKIHNENKGLRLRAREGVKANSPIIMFDDPDYLBELNACEE

CIKILDEQEDVLVFNFAIVESKNVISYKGFDFNSGFYSKSEFVKIANKNLVPTMWG

KLIRKLYLEAFASLRLEKDVKNNAEDVLVYPMLSQAQIAYMNCNLHYVYVNNNS

ICNTWNEVLKNNIQELQLVLNLRQNYILNKYCSVLVLIKLYLYIQIYKIKRTKLM

VTLAKINILTKILFKTKFLKQC"

6038..6913

/gene="cstII"

6038..6913

/gene="cstII"

/codon\_start=1

/transl\_table=11

/product="alpha-2,3-/alpha-2,8-sialyltransferase"

/protein\_id="AAK96001.1"

/db\_xref="GI:15430483"

/translation="WKVLIISGNSLSKEIDYSRLPNDPDRFCNQYFEDKYVLGKK

CKAVYFNPSPLEFEQYTYLKLHIQNOEYETELIMSNFQAHLNENQVFTFYDYPDA

HLGYDFFKQLKEFNAYFKHEIYNQRTISGVMCTVAIALGKEIYLSGIDFYQNGS

SYAFDTQKNLLKLAPNFDNNSHYIGHSKNTDKALDFLEKTYIKYLCPLNSLLA

NFIELAPNLNSNFIQEKNNYTKDILIPSSAYGKFTKNINFPKIKIKENIYYKLIKID

LLRLPSDIKHYFKGK"

6914..7954

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS	/gene="neuB1"		Db	6158	CTTGGTAAAAATGCAAGCAGTATTTTACAATCCTAGTCTTTTTTTTGGTGAACAATACTAC	6217
	6914..7954		Qy	61	ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn	80
	/codon_start=1		Db	6218	ACTTTAAACATTTAATCCAAATCAAGATATGAGACGAACTAATCATGTGTTCTAAT	6277
	/transl_table=11		Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAAsnPheValLysThrPheTyrAspTyrPhePro	100
	/product="putative sialic acid synthase"		Db	6278	TTTAACCAAGCTCATCTAGAAATCAAAATTTTGTAAAAACTTTTACGATATTATTTCTCT	6337
gene	/db_xref="GI:15430484"		Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
	/translation="MKETKIQNIISBEKAPLVPEIGINHGSLDELAKIMVDAFSA		Db	6338	GATGCTCATTTGGGATATGATTTTTCACCAACACTTAAAGAAATTCATATGCTTATTTAAA	6397
	GAKILKHOTHIVEDMSKAAKVIIPGNAKISITETMOKCALNYKDELALKEVTEKLG		Qy	121	PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
	VLTSPSRAGANLEDMGVSAFKIGSGECNNYPLIKHIAAFKPMIVSTGNLSLESI		Db	6398	TTTCAGCAAAATTTATTTCAATCAAGAAATTTACCTCAGGGGTCTATATGTGCACAGTAGCC	6457
	KPTVKILDLNDIEPVLMTNLNYPHNLVRLNMLKLGKGFSCVMGLSDHTDNLAC		Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
CDS	/db_xref="GI:15430485"		Db	6458	ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATGATTTTATCAAAATGGATCA	6517
	/translation="MKKILFITGTRADYSKIKSLMYRVONSSEFELYIFATGMHLSKN		Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn	180
	FGYTVKELYKNGFNIVFEINVDKYOTDKALATTIDGFSRYVNEFKDPLIVHGDRI		Db	6518	TCTTATGCTTTTCATACCAACAAACAAAAATCTTTTAAAAATTTGGCTCCTAATTTTAAAAAT	6577
	EPLAAVVGALNVLVHIEGGEISGTIDDSLRAISKLAIHLVNDFAKRLMLQIG		Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
	EDKSLITIGSPDLVNDKISISEAKYIDYINKNYALLMFHPVTTEITSIKNQAD		Db	6578	GATAATTCACATATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTA	6637
gene	/db_xref="GI:15430486"		Qy	201	GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
	/translation="MSLAIIPARGSGKIKNNKLVLLNNKPLIYYTIKAALNTRKSIX		Db	6638	GAATAAACTAGCAATAAAGCTATATTTGTTATGTCCTAACAGCTCTTTAGCAAAATTTT	6697
	VWVSDSDIEILNYSQNDVILKRPISLAQDDTTSDKVLHLAKFYKDKVDFLOFT		Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAspTyrThr	240
	SPLNTIHDIAFNLKNSNANALISVSECNKKILKAFVCDNDYDGLAGICNDVEPFI		Db	6698	ATAGAACTAGCGCAAAATTTAAATTTCAATTTTATCATACAGAAAAATAACTATACT	6757
	RQKLPTYMNGAIYILKIEFLNPSFLQNKTHFLMDESSLDIDCLEDLKVEQI		Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
CDS	/db_xref="GI:15430487"		Db	6758	AAAGATATATCTCATCTCTTAGTGAGGCTTATGAAAAATTTTCAAAAAATATTAAATTTT	6817
	/translation="MSLAIIPARGSGKIKNNKLVLLNNKPLIYYTIKAALNTRKSIX		Qy	261	LysLysIleLysLysGluAsnValTyrTyrLysLeuLysLeuLysAspLeuLeuArgLeu	280
	SPLNTIHDIAFNLKNSNANALISVSECNKKILKAFVCDNDYDGLAGICNDVEPFI		Db	6818	AAAAAAATAAAAATTAAGAAAAATTTATTATCAAGTTGATAAAGATCTATTAGATTAA	6877
	RQKLPTYMNGAIYILKIEFLNPSFLQNKTHFLMDESSLDIDCLEDLKVEQI		Qy	281	ProSerAspIleLysHisTyrPheLysGlyLys	291
	/note="orf11; premature translation stop due to deletion		Db	6878	CCTAGTGATATAAGCAATTTTTCAAAGGAAAA	6910
Alignment Scores:	Pred. No.: 3,43e-182		Length: 11455		AF401529 12388 bp DNA linear BCT 03-JAN-2002	
	Score: 1511.00		Matches: 278		Campylobacter jejuni O:23 heptosyltransferase I (waaC) gene,	
	Percent Similarity: 98.6%		Conservative: 9		partial cds; lipid A biosynthesis acyltransferase (htrB), putative	
	Best Local Similarity: 95.5%		Mismatch: 4		two-domain glycosyltransferase, putative glycosyltransferase,	
	Query Match: 96.7%		Indels: 0		truncated beta-1,4-N-acetylglucosaminyltransferase (cgtA-I),	
US-10-734-719-9 (1-291) x AY044868 (1-11455)	DB: 15		Gaps: 0		truncated beta-1,3-galactosyltransferase (cgtB), bifunctional	
	1		2		alpha-2,3/-2,8-sialyltransferase (cst-II), sialic acid synthase	
	3		4		(neuB1), putative N-acetylglucosamine-6-phosphate 2-epimerase	
	5		6		(neuC1), beta-1,4-N-acetylglucosaminyltransferase (cgtA-II),	
	7		8		CMP-Neu5Ac synthetase (neuA1), putative acetyltransferase, and	
Qy	1		2		putative glycosyltransferase (waaV) genes, complete cds; and	
	3		4		heptosyltransferase II (waaF) gene, partial cds.	
	5		6		AF401529	
	7		8		AF401529.1 GI:15718497	
	9		10		Campylobacter jejuni	
Db	11		12		Campylobacter jejuni	
	13		14		Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;	
	15		16		Campylobacteraceae; Campylobacter.	
	17		18		1 (bases 1 to 12388)	
	19		20		Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E.,	
Qy	21		22		LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	
	23		24		CTACCAATGATTGTTGATGTTTATGATGCAATCAATTTTATTTTGAAGATAAATCTAT	
	25		26		LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	
	27		28		1	
	29		30		1	



```

FGYTVKELYKNGFNRYBYFVINYQYQTDKALATTIDGFSRYVYNELKPDLIWVHGDR1
BPLAAAVIGALNLIWAHIEGGEISGTTDDSLRHAISKLAIHLVNDVFAKRLMOLG
EDGKSIIFIIGSPDLLELNNKISLNEAKYVDINVENVALIMFHPVTTETISIKNOAD
NLVKALQSNKYVIVYNNDLGFELILQSYEELKNNPRFKLPFSLREYFIFLTKNA
DFIIGNSCILKLEALYTAGILVSRONGRLNGENTILKVNANSDEILKAINTHHKQ
DLFTKLEILDSKLFPEYLLSGEFKLSQVKFKDIXHA"
9031..9978
/gene="cgtA-II"
9031..9978
/gene="cgtA-II"
/feature="orf5-II; phase-variable gene due to homopolymeric
G-tract"
/codon_start=1
/transl_table=11
/product="beta-1,4-N-acetylgalactosaminyltransferase"
/protein_id="AAL06007.1"
/db_xref="GI:15718507"
/translation="MLKKIIISLYKRYISKKLVLDNEHFIFKKNKIYGGKHGFFDFD
EKADVKSPNPMGFIKRVKNALTLRVLSLILPALORGIITAYNDCDDGSBELILEFC
KQYNFTAKKPYKVDLENKNEENKLYSYNMAASFIPLDMEFKIDVDHYDAKLL
KSFYRIDQENKALCYPRINFIILNGNIYVONSNGYFISGGDQLIKRNSSFERR
VSKSQWDPKGLLEELYSEQVLSQGVKLOAPLLOHFFALKYRRNDYQOYLDLS
LEEFQAFHRKSKAKKIDFAMLRPVIEQLKXFGGEIK"
9975..10640
/gene="neuA1"
9975..10640
/gene="neuA1"
/codon_start=1
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAL06008.1"
/db_xref="GI:15718508"
/translation="MSLAIIPARGSGKGIKNKLVLLNNKPLIYYITIKALNAKSISK

```

## Alignment Scores:

```

Pred. No.: 3,07e-180 Length: 12388
Score: 1496.00 Matches: 275
Percent Similarity: 97.3% Conservative: 8
Best Local Similarity: 94.5% Mismatches: 8
Query Match: 95.7% Indels: 0
DB: 15 Gaps: 0

US-10-734-719-9 (1-291) x AF401529 (1-12388)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6001 ATGAAAAAGTTATTATTCTCGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGA 6060

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6061 CTACCAAAATGATTTTGTAGTATTTAGATGTAATCAATTTTATTTTGAAGATAAATACTAT 6120

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6121 CTGTGTAATAAATGCAAGACGAGTATTTACAACTCTAGTCTTTTGTGAAACAATACTAC 6180

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6181 ACTTTAAACATTTAATCCAAATCAAGATATGAGATCGAACTAATATGTTCTAAT 6240

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnGluAsnGluValPheTyrAspTyrPhePro 100
Db 6241 TACACCAAGCTCATCTAGAAAAATGAAAAATTTGTAAAAAATTTTACGATATTATTCCT 6300

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6301 GATGCTCAATTTGGGATATGATTTTTCACAAACAACTTAAAGAAATTCATGCTTTATTTAA 6360

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6361 TTTACGAAATTTATTTCAATCAAGAAATTTACCTACGGGGTCTATATGTGCACAGTAGCC 6420

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160

```

```

Db 6421 ATAGCCCTAGGATACAAAGAAATTTATCTTCTCGGAATTCATTTTATGATAATGGGGG 6480
Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn 180
Db 6481 GGGTATGCTTTTGTATACCAACAAACAAATCTTTTAAAAATTTGGCTCTCTTAATTTTAA 6540
Qy 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6541 GATAATTCACATATATCGACATAGATAAATAACAGATATAAAGCTTTAGAAATTTCTA 6600
Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6601 GAAAAAATCTACGAAATTAAGCTATATTTGTTTATGCTTCTAACAGCTCTTTAGCAAAATTT 6660
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 6661 AVAGAACTAGCAAAATTTAAATTCAAATTTTATCATACAAGAAAAATAACTATACT 6720
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 6721 AAAGATATACTCATACCTTCTAGTAGGCTTATGGAATAATTTACAAAAATATTAATTTT 6780
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 6781 AAAAAAATAAAAAATTAAGCAATTTATTTCAAGGAAATAATTTATCAAGTTGATAAAGATTAT 6840
Qy 281 ProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6841 CCTAGCGATATAAGCAATTTATTTCAAGGAAAA 6873

RESULT 67
AY422196 12370 bp DNA linear BCT 24-FEB-2005
LOCUS
DEFINITION
Campylobacter jejuni putative heptosyltransferase I gene, partial
cds: putative lipid A biosynthesis acyltransferase, putative
two-domain glycosyltransferase, putative glycosyltransferase,
putative beta-1,4-N-acetylgalactosaminyltransferase, putative
beta-1,3-galactosyltransferase, putative sialic acid synthase, putative
sialyltransferase, putative UDP-N-acetylmannosamine synthase,
UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine synthase,
putative beta-1,4-N-acetylgalactosaminyltransferase, CMP-Neu5Ac
synthetase, putative acetyltransferase, and putative
glycosyltransferase genes, complete cds; and putative
heptosyltransferase II gene, partial cds.
AY422196
AY422196.1 GI:40217881
KEYWORDS
SOURCE
Campylobacter jejuni
ORGANISM
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 12370)
Gilbert, M., Karwasaki, M.-F., Godschalk, P.C.R., Brochu, D., Endtz, H.P.
and Cunningham, A.
Sequencing of the lipooligosaccharide biosynthesis locus of
Campylobacter jejuni MF6
Unpublished
2 (bases 1 to 12370)
Gilbert, M.
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
Location/Qualifiers
1..12370
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="MF6"
/serotype="HS:4 / HS:64"
/db_xref="taxon:197"
<1..325
/codon_start=2.
CDS

```



```
CDS
9980..110645
/codon_start=11
/transl_table=11
/product="CMP-Neu5Ac synthetase"
/protein_id="AAR82861.1"
/db_xref="GI:40217891"
/translation="MSLAIIIPARGSGKGIKNKLVLLNNKPIYYITIKAAALNAKISK
VLVSSDSVILNYAKSONVDILKRPISLAQDDTSDKVLHAKFKYKQYEDVIFLOPT
SPLRNIHDKAFNLKYNANALISVECONKILKAFVNDYDGLAGICNDYEPMP
ROKUPKTYMSNGAIYILKIFELNPNPSLQNKTKYFLMDESSLDICLEDLKKVEQI
WKK"
10633..11463
/translation dependent on homopolymetric G-tract
polymorphism"

Alignment Scores:
Pred. No.: 4.65e-175 Length: 12370
Score: 1455.50 Matches: 269
Percent Similarity: 96.6% Conservative: 13
Best Local Similarity: 92.1% Mismatches: 9
Query Match: 93.1% Indels: 1
DB: 15 Gaps: 1

US-10-734-719-9 (1-291) x AY422196 (1-12370)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db 6003 ATGAAAAGTTATTTATTTGCTCGAATGGCCCAAGCTTAAAGAAATTCATTATTCAGG 6062

Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db 6063 CTACCAATGATTTTGATGATTTATGATGCAATCAATTTATTTTGAAGTAAATCTAT 6122

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 6123 CTGTGTAATAATGCAAGCAGTATTTTACAATCTAGCTTTTTCGAACAATACTAC 6182

Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db 6183 ACTTTAAACATTTAATCAAAATCAAGAATACGACAGAGAATTAATCGTGTCTAAT 6242

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 6243 TTTAATCTAACTCATATAGAAAGCGAAATTTTAAAGAACTTTTATGATTTATTTCT 6302

Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 6303 GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGATTTAATGCTTATTTTAA 6362

Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 6363 TTTACGAAATTTATTTCAATCAAGAATTTACCTCAGGAATCTATATGTGTCAGTAGCC 6422

Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 6423 ATAGCCCTAGGATACAAAGAAATTTATCTTCGGGAATTCATTTTATCAAAATGATCA 6482

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180
Db 6483 TCTTATGCTTTTGATCTAATAACAAATAATCTTTTAAATTCGTTCTTAATTTTAAAT 6542

Qy 181 AspArgSerHisTyrIleGlyHisSerTyrAsnThrAspIleLysAlaLeuGluPheLeu 200
Db 6543 GATAATTCACACTATATCGGACACAGCAAAATATACAGACCTTAAAGCTTTAGAAATTC 6602

Qy 201 GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220
Db 6603 GAAAAAATCTTACAAATAATAAATATATATGCTTATGCTTAAACAGTCTTTTACAAATTT 6662

Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIle--GlnGluLysAsnAsnTyr 239
Db 6663 ATAGAACTAGCCCAAAATTTAAATTTCAAAATTTTATCATAAAAAATAAATAAATACTAC 6722
```

```
Qy 240 ThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsn 259
Db 6723 ACTAAAGATATACTCATACCTTCTAGTGAGCTTATGAAAAATTTTCAAAAAATATTATT 6782

Qy 260 PheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArg 279
Db 6783 TTTAAAAAATAAAAAATTAAGAAAAATATTATTATTAACAAGTTGATAAAGATCTTATAAGA 6842

Qy 280 LeuProSerAspIleLysHisTyrPheLysGlyLys 291
Db 6843 TTACCTAGTAGATATAAGCAATTTTCAAGGAAAA 6878

RESULT 68
AY854153 12403 bp DNA linear BCT 31-DEC-2005
Campylobacter jejuni strain GB5 lipooligosaccharide biosynthesis
locus, partial sequence.
ACCESSION AY854153
VERSION
KEYWORDS
SOURCE
ORGANISM
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 12403)
Gilbert M., Brochu, D. and Karwaski, M.-F.
Sequencing of the Campylobacter jejuni GB5 LOS locus
Unpublished
REFERENCE
2 (bases 1 to 12403)
Gilbert M., Masotti, M. and Karwaski, M.-F.
Direct Submission
AUTHORS
TITLE
JOURNAL
Submitted (13-DEC-2004) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
Location/Qualifiers
source
1..12403
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="GB5"
/serotype="HS:4/64"
/db_xref="taxon:197"
misc_feature
<1..>12403
/notes="lipooligosaccharide biosynthesis locus"
CDS
<1..353
/codon_start=3
/transl_table=11
/product="putative heptosyltransferase I"
/protein_id="AAK45333.1"
/db_xref="GI:61387362"
/translations="EPAKEVLNLSGIDETHIELAPKESLELMMAFTKMDLIIGNDSG
PThLAPALNKASITIFGATPSYRNAFQTHINKIIDAGKTIQNAKHIDKSDFCITRIE
EDIFKLAKGLLNEK"
343..1230
/codon_start=1
/transl_table=11
/product="putative lipid A biosynthesis acyltransferase"
/protein_id="AAK45334.1"
/db_xref="GI:61387363"
/translation="MKNSDRIYLSYLYLFFVTFMPECILHFLIALIVARITYFLNKI
HKIIDINIQICFPQYTKERDKLSKIYENFAQFGIDCLQNTTKKELINKYNFIN
ENLLIDALKRPIIETTAHYGNWEILSLAAYKGAIAIVGKLIKSEIMVEILSOSR
TQFDIELDKKGIQWLSALKKGRILGILTDQCAENESVRLKFFNKVNYQAGSL
IAQRNALIIPVYAYKDGKFCIEFFKAKDSQNASLEELTYQAQSCSEMIKKRPWEY
FFPHRRFASYNKEIYKDAK"
1227..2774
/codon_start=1
/transl_table=11
/product="putative two-domain glucosyltransferase"
/protein_id="AAK45335.1"
/db_xref="GI:61387364"
/translation="MNLKOISVIIIVKNKQTLLECLNSLKDQDEIILLNNSSDNTL
NTANERKDFANLHIYSAFIFGALKNLALSVAKNMDWILSIDADEVLENSLKN
LELQEDNIIALSRLNLYKGEWIKACGWPDYVWRIFPKNTRFRDNLVHESLVLPSNV
```



KKTHLKNGLKHHAFRDISHLINKHMOHYSSLSWAKQNIHKKSGVLKANLRAFWTPPRNYP  
LKNGLYXGFIISVCNALGTFPKWMLYELQKQPKTCALLIITTNOKERLKLVD  
SVNRLPLPNEVLIDGMSKEDTARLIEEYQKQFPCLPKHWHQEDSEFLNKSXRNKII  
KNADSEYIIVIDGMILEKDFIHEHLEFAQRURFLQSGRVILNEKESEELKDDYRI  
AFNKQDFKSKNSFLAKMLKDEKIFKKTTLIKGIRGCNMSFFKTDPFKLDGPN  
ENFIGWREDSSEFARFLNKGIFRRLKRAIAYHYHEENSKMLESNHQIYLDITIK  
NKKISWF"

CDS

2778..3971

/codon\_start=1  
/transl\_table=11  
/product="putative glycosyltransferase"  
/protein\_id="AAK45336.1"  
/db\_xref="GI:61387365"

/translation="MKKIGVPIYVVEKVLRECLDSVINQTYTNLEIMLVNDSGTDB  
NSNIAKTYTKDKRITLFDKNGGLSSARNVGIEYPSGKYKANQTOHKHKESEELIEP  
OLDNNPNYTIKAYKSAQAFNEKDLNFTYPIIIFLDSNWEINCEIECVIRM  
KNVDLWFDHDCYENIKNGKTRMEIINFKEECIITPKYANRALSVGSRDISFG  
WNGMIDFNFLQKLFNFIINEDIHFGIILPASANKIYVLQKLYLCRLRANSIN  
HDKKITKANYSEYFKDIYETFGENAKKQVLAASRIITALKLIEFFKQKSNALA  
IKETFLPCYAKKALMIKFKKQDPLNKEQLVLKPFQITQKLPYDINKEWQKIKKIFNN  
KNKN"

CDS

4020..5078

/codon\_start=1  
/transl\_table=11  
/product="putative  
beta-1,4-N-acetylgalactosaminyltransferase"  
/protein\_id="AAK45337.1"  
/db\_xref="GI:61387366"

/translation="MLFQSYFVKIICLPIPRKIRHKIKKTFLLKNIQORDKIDSYLPK  
KTLAQKNYNEDLIKNAKAIIGBEHGFYFNDEKSKDPKSPLNPMWAFIRVNSEAITL  
KASLESLPALQGVLYNDCTDGSERILEFCQKQPSFPIKYPVEIQIONPKSEEN  
KLGSYNYVASFPKDEWLKIDVDHYDAKLYKSPYIPRKNYHVHYSIRIDIFNE  
EKFYVQNGKEGELKAPGDCIAQTNLFWKEILIEDDTFKWNTAKNNIENAKSYEIL  
KVRNRIYFTFLNNYHFPFIKNYRKNDYKQLNMVSLDDFIKNYKEKLNQIDFQWLEY  
KTLKVVYKLTSSSTSKI"

CDS

complement (5730..5942)

/note="inactive protein"  
/codon\_start=1  
/transl\_table=11  
/product="truncated beta-1,3-galactosyltransferase"  
/protein\_id="AAK45338.1"  
/db\_xref="GI:61387367"

/translation="MPKISILPFTFVVEKVIKALESCINQSPKDIIEIIVVDDCGSDK  
SIDIAKEYAKKDDRIKIIHNEENLNF"

CDS

6031..6822

/codon\_start=1  
/transl\_table=11  
/product="putative alpha-2,3/8-sialyltransferase"  
/protein\_id="AAK45339.1"  
/db\_xref="GI:61387368"

/translation="MKKVIIAGNGPSLKEIDYGRLPNDPVRNQNQYFEDKYVLGKK  
CKAVFNPSLFFQYTYLKHILQOEYETELIVCSNFNLTHTIESEFLKNFYDFPDA  
HLGVDFPKQKFNAYKPEHIVPNQRTSGIYMCVAIALGYKEIYLSGIDIFYONGS  
SYAPDTQKNLLKLVSNFKNDSHYIGHSKNTDLKALELEKTYKIKLYCLCPNSLLA  
NFIELAPLNSNFIKKKJITTKIYSYLLVRLMENQKILFKK"

variation

6741

/product="results in truncated putative alpha-2,3/8  
sialyltransferase"  
/replace=""

variation

6742^6743

/product="results in extended C-terminus of putative  
alpha-2,3/8 sialyltransferase"  
/replace="a"

CDS

6909..7949

/codon\_start=1  
/transl\_table=11  
/product="putative sialic acid synthase"  
/protein\_id="AAK45340.1"  
/db\_xref="GI:61387369"

/translation="MKEIKQNIISBEKAPLVVPEIGINHGSLLELAKIMVNAFSA  
GAKILKQTHIVDEMSKAARKVTPGNAKISIVELMOKCALDYKDELALKEYTEKGL  
VYLSTPFRAGANLEDMGVSAFKIGSGCENPNPLIKHIAAPKPMIVISGMSIESI  
KPTVKILLDNEIPFVLMTNLTNPTPHNLVRLNMLELKEKFECSMVGLSDHTTNLAC

LGAVALGACVLERHFTSDMHRSGPDIVCSMDTQALKELIIQSEQMAIMRGNNESKAA  
KQEOVTIDFAPASVUSIKDKKGEVLSMDNIWVKRPGLGGISASEFENILGKALRDI  
ENDTQLSYEDFA"

CDS

7946..9070

/codon\_start=1  
/transl\_table=11  
/product="putative UDP-N-acetylglucosamine-6-phosphate  
2-epimerase"

/protein\_id="AAK45341.1"  
/db\_xref="GI:61387370"  
/translation="MKKILFITGTRADYSKIKSLMYRVONSSEFELYIPATGMHLSKN  
FGYTVKELYKNGFNKIYEFINRYKYQTDKALATTIDGFSRYVNLKPDLIYVHGDR  
EPLAAAVIYALNLIIVAHIEGSEISGTTDDSLRHAISKLAHILHVNDFAKRLMQLG  
EDEKSIPTIGSPDLELNDNKISLNEAKKYVDINVENVALLMHFVPTTEITSKNQAD  
NLVKALQSNKNYIYIYNNDLGPELIIQSYEELKNRPFKLPSPSLREYFVITLKNA  
DPIIGNSCKITKEADITAGLIVSGRONGRIGNENTLKVANSSEILKAINTIHKKQ  
DLFSAKLEISLFFEYLOSSEFFKLSTQKVFQDIKHA"

/codon\_start=1

/transl\_table=11

/product="putative

beta-1,4-N-acetylgalactosaminyltransferase"

/protein\_id="AAK45342.1"

/db\_xref="GI:61387371"

/translation="MLKKIISLYKRYISKKLVLNDNEHFIKENKNIYEKKHKHGFOLD  
EKADVKSPLNPMGFIRVKNBALTLRVLSLESLPALQRGIIAYNDCDDGSEELILEFC  
QYFNPFIKKYPYKVDLENPKNEENKLSYNNWAASFIPLEWFIKIDVDHYDAKLL  
KSYRIDQENKALCYPRINFILNGNIYVONSNGYFIGGGDGLIKRNSFFIBRR  
VSKSQWIDPKGLIEELYSEQQVLSQVKIILQAPLLOHFPALKPYRRNDYQOYLDTLS  
LEEFQAFHRKSKEAKKIDFAMLRPVIQILKQFQGEIK"

9668

/product="results in truncated putative

beta-1,4-N-acetylgalactosaminyltransferase"

/replace=""

/product="results in truncated putative

beta-1,4-N-acetylgalactosaminyltransferase"

/replace="g"

10007..10672

/codon\_start=1

/transl\_table=11

/product="putative CMP-Neu5Ac synthetase"

/protein\_id="AAK45343.1"

/db\_xref="GI:61387372"

/translation="MSLAIIPARGSGSKIKKNVLNKNPLIYTIKALNAKSISK  
VLSSDSVDILNYAKSQNVILKRPISLAQDDTTSKVLHLLHAKFKYQYEDVIFLOPT  
SPLRNTHIDKAFNLKNSNANALISVSECNKILKAFVNDYDGLAGICNDYEYFPM  
ROKLPKTYMSGATVILKIEFLNPNPSLQNKTKYFLMDESSSLDIDCLELKKVQEI  
WKK"

10660..11490

/codon\_start=1

/transl\_table=11

/product="putative acetyltransferase"

/protein\_id="AAK45344.1"

/db\_xref="GI:61387373"

/translation="MEKITLCKNNILNLKQYNIYTKYIENPRFRSLKTKDPTIT  
PLENNQAGGIEEYCAFKFSPILHIEIGFSFGSLPHYAKFYCSIDAGVSMF  
NFQPMRISTAGTSTYETHFINDACNHINKTEPIVNHNPSSITHLIIOQDVWIG  
KDVLLKQGITLGTGCVIGORAVTKDPPVPAIVAGIPAKIIRFRDEKTIERLLKIOW  
WKYHFAFDYDIDLNLKLNQYLLLEEKIISYSIYNPNKLYFRDILELKSKKIFNL  
"

10829

/product="results in truncated putative acetyltransferase"

/replace=""

10830^10831

/product="results in truncated putative acetyltransferase"

/replace="g"

complement (11482..12294)

Alignment Scores:

7,66e-174 Length: 12403

1446.00 Matches: 269

96.2% Conservative: 13

Best Local Similarity: 91.8% Mismatches: 9			
Query Match: 92.5% Indels: 2			
DB: 15 Gaps: 0			
US-10-734-719-9 (1-291) x AY854153 (1-12403)			
Qy	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	6031	ATGAAAAGATTATTGCTGGAAATGGCCACAGCTTAAAGAAATTTGATTATTCAGG	6090
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	6091	CTACCAATGATTTTGATGATTTAGATGCAATCAATTTATTTTGAAGATAAATACTAT	6150
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyr	60
Db	6151	CTTGTAATAAATGCAAGCAGTATTTTACAACTCTAGTCTTTTTTTTGAACAAATACTAC	6210
Qy	61	ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	6211	ACTTTAAACAATTTAATCCAAATCAAGATACGACAGAAATTAATCGTGTTCTAAT	6270
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	6271	TTTAATCTAACTCATATAGAAAGCGAAATTTTAAAAAACTTTTATGATTATTTTCT	6330
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	6331	GATGCTCATTTGGGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTATTTTAA	6390
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	6391	TTTCAGCAATTTATTTCAATCAAGAAATTTACCTCAGGAATCTATATGTGTCAGTAGCC	6450
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	6451	ATAGCCCTAGATACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGAACA	6510
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180
Db	6511	TCTTATGCTTTTGATACTAAACAAATAATCTTTTAAATTTGTTTCTAAATTTTAAAAAT	6570
Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
Db	6571	GATAATTCACACTATATCGCACACAGCAAAATATACAGACCTTAAAGCTTTTGAATTTCTA	6630
Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
Db	6631	GAAAAAATTTCAAAATNAATTAATATTTGCTTATGCTTAACTTAAAGCTTTTGAATTTTCTA	6690
Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIle--GlnGluLysAsnAsnTyrT	240
Db	6691	ATAGAATAGCGCCAAATTTAAATTTCAATTTTATCATAAAAAAATAAATAACTACA	6750
Qy	240	hrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnP	260
Db	6751	CTAAGATATATCTCATACCTTCTAGTGAGGCTTATGGAATAATTTTCAAAAAATATATT	6810
Qy	260	heLysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgL	280
Db	6811	TTAAAAAATAAATAATTAAGAAATATTTTATTAAGTTGATAAAGATCTATTAAAGAT	6870
Qy	280	euProSerAspIleLysHisTyrPheLysGlyLys 291	
Db	6871	TACCTAGTATATAAGCAATTTTCAAGGAAAA 6905	
RESULT 69			
AF195055			
LOCUS			
DEFINITION Campylobacter jejuni strain MSC57360 hypothetical protein, Siae			
(Siae), and Siae (siae) genes, complete cds.			
ACCESSION AF195055			
VERSION AF195055.1 Gi:11095586			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. .3071			
/organism="Campylobacter jejuni"			
/mol_type="genomic DNA"			
/strain="MSC57360"			
/db_xref="taxon:197"			
/note="serogroup: HS1"			
59. .943			
/note="ORF1"			
/codon_start=1			
/transl_table=11			
/product="hypothetical protein"			
/protein_id="AAG29922.1"			
/db_xref="GI:11095587"			
/translation="MSMINALVCGNGPSLKNIDYKRLPKQFDVRCNQFYFEDRYFV			
GDKVYVPFVFEQYYSKLIQNEEYNIENIVCTINLEIDGQFQFVDFNDEL			
SDAFILGHEIILKKLQFFAVIKYNEIYNRORITSGVMCATAVAGYKSIYIGDIFVQ			
DTNNIYAFDNKKLLNKTCTGPNKQFKNHSMACDLOALDYLMEKYDNIYISLNSD			
EYFKLAPDIDGSDPVSJSSKPKYINDILIPDKYAQERYGKKSRUKENLHYKLKDLIR			
LPSSIKHYLKYERANKNR"			
924. .1955			
/gene="Siae"			
924. .1955			
/gene="Siae"			
/note="Sialic acid synthetase; required for sialylation of			
LOS core"			
/codon_start=1			
/transl_table=11			
/product="Siae"			
/protein_id="AAG29923.1"			
/db_xref="GI:11095588"			
/translation="MQIKIDKLITISOKNPLIIPKGINHNGSLKIAKLVDAKRAKA			
KIIKHQTHIVDEMSQAKNVIIPGNANISYIEMOCALNKDELALKEIVYKQGLVY			
LSTPFSRAANLEOMGVSAKYKIGSGECNNYPLIKHIAQFKPMIISTGWNISIESIKP			
TVKLRDYEIFPVLLHTTNLYTPPSHLVRLQAMLELYKEFNCLYGLSDHTNNLACIG			
AIALGASVLERHFTESMDRKGDPDIVCSMDSTLUKLINQTOEWVLLRGDNNKPKKE			
QVTIDFAFASVVSIIKDKKGEILSDMNIWVRPSKGGISAKDFAILGKRAKKDKNN			
IOLTWDDEP"			
1952. .3067			
/gene="Siae"			
1952. .3067			
/gene="Siae"			
/note="N-acetylmannosamine synthetase; required for			
sialylation of LOS core"			
/codon_start=1			
/transl_table=11			
/product="Siae"			
/protein_id="AAG29924.1"			
/db_xref="GI:11095589"			
/translation="MKKIVFVSGTRADFSKIKSLMKVENSPELFIFVTGMHMSKK			
EGSTYMEIEKCGKNLYKYNHDKYQMDKALSTTDGFSKFIHEIPDLIVVHGDV			
EFAAAIVGSLNNILVAHIEGSELSCTIDESLRAHSLAHILVNDIEAKRLIQMG			
EDEKSIFFIGSPDLLEUNNTISLDEAKYDIDKFNVAIAIFIPITTEISLYKQSEE			
FYNALIKSKNTYIVYPNNDLGFLQYERLKNRERFKIFPSLREYFISLLKNN			
FTIGNSSCIIEALYINGILVGRQDGTINKTIRVNAEERKIDLEALINTSKCTN			
ITNKRLEILNSSEQFVRLKNNILFTINKQKIPMDKK"			
ORIGIN			

## Alignment Scores:

```
Pred. No.:      2,75e-92      Length:      3071
Score:          804.00      Matches:      158
Percent Similarity: 69.3%      Conservative: 45
Best Local Similarity: 53.9%      Mismatches: 72
Query Match:      51.4%      Indels:      18
DB:              15          Gaps:      6

US-10-734-719-9 (1-291) x AF195055 (1-3071)

Qy  5  IleileAaGlyAsnGlyProSerLeuLysGluileAspTyrSerArgLeuProAsnAsp 24
Db  80  TTGGTATCGCGTAATGACCTAGTTTAAAAAATATAGATTATAAAGGTTACCAAAACAA 139
Qy  25  PhaAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLys 44
Db  140  TTTGATGTATTATAGATCAATCAGTTTATTTTGAAGATAGATATTTTGGTGGGTAAGAT 199
Qy  45  CysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHis 64
Db  200  GTAAATATGTTTTTTTAACTCTTTGTGTTTTTGAACAAATCTACTAGTAAAAAA 259
Qy  65  LeuileGlnAsnGlnGluTyrGluThrGluLeuileMetCysSerAsnTyrAsnGlnAla 84
Db  260  CTTATACAAACGAAGAATAATACATAGAAATATAGTTGTTCTACTATTAAATTAGAA 319
Qy  85  HisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeu 104
Db  320  TACATAGATGGATTTCAAATTTGTTGATAAATTTTGAATTAATCTTAGTGATGCTTTT 379
Qy  105  GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle 124
Db  380  GGTGATGAGATTATAAAAGCTTAAAGATTTTTCCTATATTAATTAATTAAGAAATA 439
Qy  125  TyrPheAsnGlnArgileThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly 144
Db  440  TACAATAGGCAAGAATTAATCTAGTGAGTTATATGTTGCTACAGCTGTGGCTCTGGA 499
Qy  145  TyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer---TyrAla 163
Db  500  TATAAAAGTATATATAAGTGGTATTGATTTTATCAAGATACAAATAAATTTATATGCC 559
Qy  164  PheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSer 183
Db  560  TTCGATAATAATAAAAAAATTTATAATAAATGACCTGCTTTTAAAAAATCAAAATTC 619
Qy  184  HisTyrileGlyHisSerLysAsnThrAspileLysAlaLeuGluPheLeuGluLysThr 203
Db  620  AAATTATCAATCATTCATGCGCATGTTGATTACAGCTTTAGATTATCTAATGAAAGG 679
Qy  204  TyrLysileLysLeuTyrCysLeuPheProAsnSerLeuLeuAlaAsnPheIleGluLeu 223
Db  680  TATGATGTGAATATTATAGTTTA-----AATTC-----GATGAGTATTTAAATG 727
Qy  224  AlaProAsnLeuAsnSerAsnPheIleileGlnGluLys---AsnAsnTyrThrLysAsp 242
Db  728  GCTCCGATATTCGAAGTGGATTTTGTATTGATGATAAAGAACCAAAATATATATAATG 787
Qy  243  IleLeuileProSerSerGluAla-----TyrGlyLysPheSerLysAsnIle 258
Db  788  ATATTGATACAGTAAAGTATGATGCAAGAAAGATATTATGCA----- 829
Qy  259  AsnPheLysLysIleLysLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeu 278
Db  830  -----AAAAAAGTAGGTTAAAAAGAAATCTCCATTATAGCTAAATAAAGATCTAAT 883
Qy  279  ArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
Db  884  AGACTTCCAAAGTGACATAAAACACTACTTAAAGGAAAAA 922

RESULT 70
AF257460
LOCUS
```

## DEFINITION

Campylobacter jejuni 2,3-sialyl transferase (cstIII), putative sialic acid synthase (neuB1), N-acetylmannosamine synthetase protein (neuC1), and beta 1,4-N-acetylgalactosaminyltransferase/ CMP-NeuNAc synthetase fusion protein (cgt/neuA1 fusion) genes, complete cds.

ACCESSION AF257460  
VERSION AF257460.1 GI:11527837

## KEYWORDS

## SOURCE

## ORGANISM

Campylobacter jejuni  
Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.

## REFERENCE

1 (bases 1 to 4668)

## AUTHORS

Guerry, P., Ewing, C.P., Hickey, T.E., Prendergast, M.M. and Moran, A.P.

## TITLE

Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of Campylobacter jejuni

## JOURNAL

Infect. Immun. 68 (12), 6656-6662 (2000)

## PUBMED

## REFERENCE

2 (bases 1 to 4668)

## AUTHORS

Guerry, P. and Ewing, C.P.

## TITLE

Direct Submission

## JOURNAL

Submitted (18-APR-2000) Enteric Diseases, Naval Medical Research Center, 503 Robert Grant Ave., Silver Spring, MD 20910, USA

## FEATURES

Location/Qualifiers

1..4668

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strains="MSC 57360"

/serotype="HS01"

/db\_xref="taxon:197"

/note="type strain of HS01 serotype"

53..937

/gene="cstIII"

53..937

/gene="cstIII"

/note="Cst; involved in sialylation of LOS core"

/codon\_start=1

/transl\_table=11

/product="2,3-sialyl transferase"

/protein\_id="AAG37020.1"

/db\_xref="GI:11527838"

/translation="MSMNINALVCGNGPSLKNIDYKRLPKQDFVRCNOFYEDRYFV

GKDVLYFNPFPPEOYYSKLIQNEENITVCSINLEYIDGFOVDNDFELYF

SDALYGEHIIKKLKDFFAYIKYNEIYNQRITSGVVMCATAVAGYKSYISGIDFVQ

DTNLYAFDNNKNNLLNCKTGFKQKQKPFINHSMACDQLALDYLKRYDNIYSLNSD

BYFLAPDIGSDFVLSKRPKPKYINDILIPDKYAQERYGKSRLENLHYKLIKDLIR

LPSTIKHYLKEKYANKNR"

918..1949

/gene="neuB1"

918..1949

/gene="neuB1"

/inference="non-experimental evidence, no additional

details recorded"

/note="NeuB1"

/codon\_start=1

/transl\_table=11

/product="putative sialic acid synthase"

/protein\_id="AAG37021.1"

/db\_xref="GI:11527839"

/translation="MQIKDKLTIQKKNPLIIPGIGNHNGSLIAKLMDVDAKRAGA

KIHKOTHVSDMSOEAKNVTPGNANISIVEMQCALNYKDELAKYEVKQGLVY

LSTPFSRAANRLDMGVSAKYIGSGECNNYPLIKHIAQFKPKPMIISTGMNSIESKP

TVKILRDYEIFVLLHTNLYTPSHLVRQLAMLELYKEFNCLYGLSDHTNNLACIG

AIALGASVLERHTSMRKGPDIVCSMDESTLKDILNQTOEMVLRLGRNNKNPLKEE

QWITDFAFASVVSINKDKKGEILSDMNIWVKRPSKGGISAKDFAILGKRKKDKIKNN

IQLTWDDFE"

1946..3061

/gene="neuC1"

1946..3061

/gene="neuC1"

/note="NeuC1"

/codon\_start=1

/transl\_table=11

```

/product="N-acetyl mannosamine synthesis protein"
/protein_id="AAG37022.1"
/db_xref="GI:11527840"
/translation="MKKIVFVSGTRADFSKIKSLMKVNSNEFELFIVFTGMHMSKK
FGSTWMEKCGFNKIYKINHDKYIQMDKALUSSTIDGFSKFIHEIEPEPLIVVHGDRV
BPLAAIVGSLNNILVAHIEGELSGTIDESLRHAIKSLAHILVNDIEAKKRLIQMG
EDERSIFITGSPDLLENNITSLDEAKKYIDFKFNAYATAI FPHITTEINSYKOSSE
FVNALIKSEKNYVIYPNNDLGFELILQNYERLKNNERFKI FPSLRFEYFISLLKNAF
FLIGNSSCIIKEALYNLNGILVSRQDGTINKTIRVNAEKEIDLEAILNATSKCTN
ITNKRLEILNSSEQFRLKNNILFTINKQKIFMDKK"
3058. .4668
/gene="Cgt/NeuA1 fusion"
3058. .4668
/gene="Cgt/NeuA1 fusion"
/feature="Cgt/NeuA1; apparent fusion of beta
1,4-N-acetylglucosaminyltransferase and CMP-NeuNAc
synthetase"
/codon_start=1
/transl_table=11
/product="beta 1,4-N-acetylglucosaminyltransferase/
CMP-NeuNAc synthetase fusion protein"
/protein_id="AAG37023.1"
/db_xref="GI:11527841"
/translation="MTLFVYKIIAFLRLKIDKLFNDNEYFLNKKKIYNEKHGFFD
FDPNSKTKSPLNPAWAFIRVKNEAATLRVSLSEMLPAIORGVIGYNDCTDGSSEIILE
ECKQYPSIFPKYVPEHVOLENQSEENKLHSYNYVASFIPODEMLIKIDVDHYVDK
KLYFSYWASKNATVRPRINFLIIDLKIVIQNIGCGFIDGGDQLDIQKNSVFIERM
VSKQSQWDEPKTVKLESEQQIIPKHKILOAELQWHFPALKYHRNDYQKHLDAIT
LEDFKIHYRHKIKKINIMLDEKVIRELIDKFLKSLGKMTLAIIPARAGSKGIKNK
NLAILHRLPRLYITNAAKNSYKVDKI VLSDDGDIILEYQTQGVVDLKRPRKELADD
TTSDKVLHTLSFYKDYENI VLIQTPSLRTNVHIDEAFLKFKNSNNAISVSRCDN
KILKAFIDNGLKIDCNKYPFMPRQKLPKTYMNGALYIVKSNLFLNNPFLQKFT
SCYIMDEKASDIDITDDELKRVNNISFL"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 4,12e-92 Length: 4668
Score: 804.00 Matches: 158
Percent Similarity: 69.3% Conservative: 45
Best local Similarity: 53.9% Mismatches: 72
Query Match: 51.4% Indels: 18
DBs: 15 Gaps: 6

US-10-734-719-9 (1-291) x AF257460 (1-4668)

QY 5 llelleAlaGlyAsnGlyProSerLeuLysGluilleAspTyrSerArgLeuProAsnAsp 24
   |||||
   |||||
   |||||
Db 74 TTGGTATGCGGTATGGACCTAGTTTAAAAAATATAGATTATAAAGGTACCAAAACAA 133

QY 25 PheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLys 44
   |||||
   |||||
   |||||
Db 134 TTGTGATGATTTAGATGCAATCAGTTTATTTTGAAGATAGATATTTTGTGGGTAAAGAT 193

QY 45 CysLysAlaValPheTyrThrProsnPhePheGluGlnTyrTyrThrLeuLysHis 64
   |||||
   |||||
   |||||
Db 194 GTAAATATGTTTTTTTAAATCCCTTTGTGTTTTTGTGAACAATACCTATATAGTAAAAA 253

QY 65 leuilleGlnAsnGlnGluTyrGluThrGluLeuilleMetCysSerAsnTyrAsnGlnAla 84
   |||||
   |||||
   |||||
Db 254 CTTATACAAACGAGAAATATACATAGAAATATAGTTTGTCTACTATTAATTAGAA 313

QY 85 HisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeu 104
   |||||
   |||||
   |||||
Db 314 TACATAGATGATTTCAATTTTGTGATAATTTTGAATTTATCTTTAGTCATGCTTTTTPA 373

QY 105 GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluille 124
   |||||
   |||||
   |||||
Db 374 GGTCTAGATATATAAAAGAGCTTAAAGATTTTTTTTGGCTATATATAAATATATAGTAATA 433

QY 125 TyrPheAsnGlnAlaGlyIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly 144
   |||||
   |||||
   |||||
Db 434 TACAATAGGCAAGAAATTTACTAGTGGAGTTTATATGTGTGTACAGCTGTGGCTCTGGGA 493

```

```

QY 145 TyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer---TyrAla 163
   |||||
   |||||
   |||||
Db 494 TATAAAAGTATATATATAAGTGGTATTGATTTTATCAAGATACAAATAAATTTATATGTC 553

QY 164 PheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSer 183
   |||||
   |||||
   |||||
Db 554 TTGCATAATAATAAAAAAAATTTATTAAATAAATGCATGGTGGTTTAAAAAATCAAAAATTC 613

QY 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLysGluLysThr 203
   |||||
   |||||
   |||||
Db 614 AAATTTATCAATCAATCAATGGCATGTAITTTACAAGCTTTAGATTATCTAATGAAAGG 673

QY 204 TyrLysIleLysLeuTyrCysLysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223
   |||||
   |||||
   |||||
Db 674 TATGATGTGATATTATTATAGTTTA-----AATTCA-----GATGACTATTTTAAATTG 721

QY 224 AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys---AsnAsnTyrThrLysAsp 242
   |||||
   |||||
   |||||
Db 722 GCTCCGATATTGGAAGTGATTTTGTATTGATGATAAAACCCAAAAAATATATAAATGAT 781

QY 243 IleLeuIleProSerSerGluAla-----TyrGlyLysPheSerLysAsnIle 258
   |||||
   |||||
   |||||
Db 782 ATATTGATACACAGATAAGTATGCACAAGAAAGATATTATGGA-----823

QY 259 AsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeu 278
   |||||
   |||||
   |||||
Db 824 -----AAAAAAGTAGTTTAAAGAAAATCTCCATTATAAGCTAATAAAGATCTAATT 877

QY 279 ArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
   |||||
   |||||
   |||||
Db 878 AGACTTCCAAGTGACATAAAACACTACTTAAAGGAAAAA 916

RESULT 71
AF400047 13484 bp DNA linear BCT 03-JAN-2002
LOCUS
DEFINITION
Campylobacter jejuni heptosyltransferase I (waaC) gene, partial
cds; lipid A biosynthesis acyltransferase (htrB), putative
two-domain glycosyltransferase, putative glycosyltransferase,
putative galactosyltransferase, beta-1,3-galactosyltransferase,
alpha-2,3-sialyltransferase (cst-III), sialic acid synthase
(neuB1), putative N-acetylglucosamine-6-phosphate 2-epimerase
(neuC1), beta-1,4-N-acetylglucosaminyltransferase/CMP-NeuAc
synthetase fusion protein (neuA1), and putative glycosyltransferase
genes, complete cds; heptosyltransferase II (waaF) gene, partial
cds; and unknown genes.
AF400047
AF400047.1 GI:15150361
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
REFERENCE
1 (bases 1 to 13484)
Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E.,
Michniewicz,J., Cunningham,A.M. and Wakarchuk,W.W.
The Genetic Bases for the Variation in the Lipo-oligosaccharide of
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF
SIALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
J. Biol. Chem. 277 (1), 327-337 (2002)
11689567
2 (bases 1 to 13484)
Gilbert,M., Michniewicz,J., Karwaski,M.-F., Cunningham,A. and
Wakarchuk,W.W.
Direct Submission
Submitted (14-JUL-2001) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A
0R6, Canada
FEATURES
Location/Qualifiers
1. .13484
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43430"
/serotype="0:2"

```

```
/db xref="ATCC:43430"
/db xref="taxon:197"
/notes="Penner-type strain"
<1..348
/genes="waaC"
<1..348
/genes="waaC"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAK85412.1"
/db xref="GI:15150362"
/transl_table="FAKEINLSGIDETHIELAPKFNLEBIAMFTKMNNLIIGNDSGP
THIAFALNKASITIFGATPSYRNAFOQHINKIDAGKKIQNAKHIDKSDFCITRIEBE
DIFKLAKGLLNEK"
338..1225
/genes="htrB"
338..1225
/genes="htrB"
/codon_start=1
/transl_table=11
/product="lipid A biosynthesis acyltransferase"
/protein_id="AAK85413.1"
/db xref="GI:15150363"
/transl_table="MKNSDRIYLSLYILKFFVTFMPECILHFLALIVARIAFYLNKK
HRKINTNLOICPQYQKERDKLSKIYENEAQFGDCLONQNTTKEKILNKVNFN
ENFLIDALNRRIIFTHAYGNWEILSLAYAAKYGAISVGKKLKSEINWEILSQSR
TQDIEDLKGIRQMSALKRGKRTIGLITDQDCVENESVRLKFFNKEVNYQNGASL
IAQRSNALITPVAYKEDGKFCIEFFKVKDSQNASLEBELTYQAQSCCEMIKKRPWEY
FFFHRRPASYNKBIYDAK"
1222..2769
/notes="orf3"
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAK85414.1"
/db xref="GI:15150364"
/transl_table="MNKIQISALIIYQNAQTLECLNSLKDPDEIILLANESSDNTL
NIANFKOPANLHIYHSAFIGFALKKNLALSAYKNDWILSDDADEVNECIELKLN
LEIQEDNIIALSKNLKYGEWIKACGWWPDYVWRIFNKNFTFRFNDNLVHESLVLPNA
KKIYLNGLRHAFRDIIDSLIDKQYYSYLWAKQNIHKSGVLKANLRAFWTFRNFT
LKNGLFYGYGFIISVCNALGTFFPKMKLYSLOKQKPTCALIITTYNOKERLKLVD
SVKNLAPLNEVLIAADGSKEDTARLIEEYQKPPCPLKHIWQDEGFRAAKSRNNAI
KAKSEYIILLDGMILEKDPVKNHLEFARKVILQGSRVILKNSEEEILLKNNYSL
AFNKKGFKNQRSIFLAKCVYKFSKLTTKKFFKKSQLVKSGSKTNNKSFYKSDFAEAGEFN
EKTIGMGREDSFVAREFLFNDGVFRLKFNALAYHIYHEENSKNNLEINHOIYLETIK
NEKTTWR"
2773..3945
/notes="orf4"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAK85415.1"
/db xref="GI:15150365"
/transl_table="MKVGVVPIYNNVEKYLKECLDSVINQTYTNLEIILVNDGSTDB
NSLNAKEYTLKDKRIILFDKNGGLSARNVGIEYFSGEYKLNKTOITKENSLSLEP
GDPNQNTIYKYNALASDKLEISKIDYIIFLDSDNWELNCIEECVVRMNQVVDLW
FDHACIYDDGIDEGKQRMNVNFTQECTITPRDYAKQAIKVGRSDISFSWGGMIDF
SFLQKLKTKINKIINEDIHFGVNLFAADSIIYILPKRLYLCLRANSISNHDKKVTK
ANVSEYFKDLYEPFEGENAKKAYLKAASRMITALELIEFPKQKNENSOAIKEAFLP
FYVKALMIKKFKKQPLNLEKPLIIPKFIQTQTKPYDLMKIWKQIKGLDKINPAK"
complement(3929..4924)
/notes="orf5"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK85416.1"
/db xref="GI:15150366"
/transl_table="MQIQNNSLIYNTLTKKLSFPIPIKSTRRLRNHIQYKLBBPKV
TMNLSNNYINPFLEGIKPIPHDFEKKHYFKNDKIWQFWQCKNQASPMIQCCFNYSVG
QKDDYTIITLDKNIKDYLDFPPEVIEKLENNFGEKTIITFPSDLRLVCLLAYGGI
WCDAISIFLSKSIPELCKDQFPFERSKIKPSKELKKFIKSGYFSYGYNWDDDFKV
KMUSSEFMIAKNTNTFIOTLKDILMNYWNQESSENHYFYLIHIFELLKEHGFVNDIY
```

```
CDS
4982..6151
/notes="orf6"
/codon_start=1
/transl_table=11
/product="putative galactosyltransferase"
/protein_id="AAK85417.1"
/db xref="GI:15150367"
/transl_table="MKTGVGVPIYNNVEKYLKECLDSVINQTYTNLEIILVNDGSTDE
NSLNAKEYTLKDKRIILFDKNGGLSARNVGIEYFSGEYKLNKTOITKENSLSLEP
NIEGNPEYIYTVKSYKAFNNEQDLNFTPIIDYIIFLDSDDWELNCIEECVPRM
DGVDFWFDSEIYHDKIEKSYFKHSHLDKINIKRECRINPIEWLKLRONKLIKDAFA
WSGIDIPDYIKQKMKDAIFAEHDLFGIILLFSQAKNIYVYKPFYFYRIRANSJTN
QDKKITDNILPYKDFPIAEPENATLAKFYKVSWEVTSLELVRFVENYHDKISS
LKDITILFYIKNAFKIKPKDKPLCIKEKLQLLKPYSKIRYKVMKELISLRN"
complement(6711..7052)
/notes="premature stop codon due to deletion in
homopolymERIC tract"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAK85418.1"
/db xref="GI:15150368"
/transl_table="MSQISIILPTYNVEKYIARALESCTINQTFKDIIEIIVDDCGNDK
SIDIAKYASKDDRIKIHNENLKLRAVEGAKVATSPYIMFLDSDDYELNACEE
CIKILDMGGGG"
7107..7991
/genes="cst-III"
7107..7991
/genes="cst-III"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAK85419.1"
/db xref="GI:15150369"
/transl_table="MSMNINALVCGNGPSLKNIDYKRLPKQDFVRCNFQPEDRYFV
GKDVYVFNFPVFVEQYVTSKLIQNEENIENIVCSTINLEYIDGFOVNFELFY
SDAFLGHEIKKLDFFAYIKYNEIYNRQITSGVYMCATAVALGKSYIIGSIDFYQ
DTNNLAFADNNKNLAKCTGCKNOKKFINHSMACDLQALDYLAKRYDNNIYLSND
EYFKLAPDIGSDFVLSKPKKIINDILLIPDKYAQERYGKSKRLKENLHYKLIKDLIR
LPSDITKILKEKYANKNR"
7972..9003
/genes="neuB1"
7972..9003
/genes="neuB1"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAK85420.1"
/db xref="GI:15150370"
/transl_table="MQIKLDKLTISQKNPLIPIEIGINHSLSLEIAKLMVDAAKRAGA
KILKHQTHIVEDEMSQEAKNVPGNANISYIEMEQQCALNYKDEALKYVEKQGLVY
LSTPFSRAANRLDEMGVSAYKISGECNNYPLIKHIAQFKPMIISTGMNSESIRK
TVKILRDEIPIFVLLHTNTPSHLVRLQALMELKFEKCYLGLSDRTNNLACIG
AIALGASVLERHFTDTMDRGPDIVCSMDESTLDKLIINOTQEMVLLRGNNKNPLKEE
OVTIDFAPASVVSIKDIKKGELSLSDNITVWRKPSKGGISAKDFAILGRAKKDIKN
IQLTWDDFE"
9000..10115
/genes="neuCl"
9000..10115
/genes="neuCl"
/codon_start=1
/transl_table=11
/product="putative N-acetylglucosamine-6-phosphate
2-epimerase"
/protein_id="AAK85421.1"
/db xref="GI:15150371"
/transl_table="MKKILVFSVGTGRADFSKISLMMKYVENSNEFELFIVTGMHMSKK
FGSTYMEIEKCGFNKIYIINHDKIYQMDKALSSITDGFSKFHEIEPDLIVHGRV
EPLAAAVIYGLNNILVAHIEGELSSTGTTDESLHAIKLAHILHVNDEIAKKRLIQMG
EDEKSIITIGSPDLLENNTISLDEAKKYDIKFNKYAIAIHPHITTEINSLYKQSEE
FVNALIKSEKNIYIYPPNDLGFELILQYERLKNNERPKIPPSLRFYFISLLKNAN
PIIGNSSCIKEALYLNINGILVGRQDRTDINKTIRVNAEKKDILEALINTSKCTN
```

```
ITNKRLEINSEQFYRLLKNNILFTINKQKIFMDKX
10112..11722
/gene="neuA1"
10112..11722
/gene="neuA1"
/codon_start=1
/transl_table=1
/product="beta-1,
4-N-acetylglactosaminyltransferase/CMP-Neu5Ac synthetase
fusion protein"
/protein_id="AAK85422.1"
/db_xref="GI:15150372"
/translation="WTLFYKIIAFLRLIKDKLKFNEYFLNKKIYDEKHGKFFD
FDPNSKDFKSLNWPWFVRKNEATTLVSLSEMLPAIQRGVIGVNDYDGDSEIILE
FCKQYPSIPVKYIPEHVQIENPQSENKLHSYNYVASFIQDEWLLIKIDVDHYDAK

Alignment Scores:
Pred. No.: 1.15e-91 Length: 13484
Score: 804.00 Matches: 158
Percent Similarity: 59.3% Conservative: 45
Best Local Similarity: 53.9% Mismatches: 72
Query Match: 51.4% Indels: 18
DB: 15 Gaps: 6

US-10-734-719-9 (1-291) x AF400047 (1-13484)

QY 5 IlelleAlaGlyAsnGlyProSerLeuLysGluLleAspTyrSerArgLeuProAsnAsp 24
Db 7128 TTGGTATCGGTAATGGACCTAGTTTAAAAAATATAGATTATATAAAGGTTACCAAAACAA 7187

QY 25 PheAspValPheAtcCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLys 44
Db 7188 TTTGATGATTTAGATGCAATCAGTTTATTTTGAAGATAGATATTTTGTGGGTAAAGAT 7247

QY 45 CysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHis 64
Db 7248 GTAAATATGTTTTTTTAACTTTTGTGTTTTTGAACATATCTACTAGTAAAAAA 7307

QY 65 LeulleGlnAsnGlnGluTyrGluThrGluLeulleMetCysSerAsnTyrAsnGlnAla 84
Db 7308 CTTATACAAAACGAAGAATATAACATAGAAATATAGTTTGTCTACTATTATTAATAGAA 7367

QY 85 HisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeu 104
Db 7368 TACATAGATGCAATTTCAATTTTGTGATAATTTTGAATTTATATCTTAGTCATCTTTTFA 7427

QY 105 GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeu 124
Db 7428 GGTCAATGAGATTATAAAGCTTAAGATTTTTTTTGGCTATATTAAATATATGAATA 7487

QY 125 TyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly 144
Db 7488 TACAATAGCAAGAAGATTACTAGTGGAGTTATATGTTGCTACAGCTGTGCTCTGGGA 7547

QY 145 TyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer---TyrAla 163
Db 7548 TATAAAGTATATATATAGTGTTATTTTATCAAGATCAAAATAAATTTATATGCC 7607

QY 164 PheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSer 183
Db 7608 TTCGATAATAATAAAAAAAATTTATAATAATGACCTGGTTTATAAATCAAAATTC 7667

QY 184 HistTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGlyLysThr 203
Db 7668 AAATTTATCAATCAATCAATGCATGTGATTTTACAAGCTTTAGATTATCTAATGAAAGG 7727

QY 204 TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu 223
Db 7728 TATGATGTGAATATTATATAGTTTA-----AATTCA-----GATCAGTATTTTAAATG 7775

QY 224 AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys---AsnAsnTyrThrLysAsp 242
Db 7776 GCTCCGGATATTGGGAAGTGATTTTGTATGTAGTAAAAAACCAAAAAATATATAATGAT 7835
```

```
QY 243 IleLeulleProSerSerGluAla-----TyrGlyLysPheSerLysAsnIle 258
Db 7836 ATATTGATACCATGATAGTATGCAAGAAGATATTATGGA-----7877

QY 259 AsnPheLysLysLysLysLysLysGluAsnValTyrTyrLysLeulleLysAspLeuLeu 278
Db 7878 -----AAAAAAGTAGGTTAAAAAGAAATCTCCATTATAGCTAATAAAAAAGATCTAATT 7931

QY 279 ArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
Db 7932 AGACTTCCAAGTGACATAAAACACTACTTAAAGGAAAAA 7970

RESULT 72
AY044156
LOCUS
DEFINITION
AY044156 13484 bp DNA linear BCT 03-JAN-2002
Campylobacter jejuni heptosyltransferase I (waaC) gene, partial
cds; lipid A biosynthesis acyltransferase (htrB), putative
two-domain glycosyltransferase, putative glycosyltransferase,
putative galactosyltransferase, beta-1,3-galactosyltransferase,
alpha-2,3-sialyltransferase (cst-III), sialic acid synthase
(neuB1), putative N-acetylglucosamine-6-phosphate 2-epimerase
(neuC1), beta-1,4-N-acetylglucosaminyltransferase/CMP-Neu5Ac
synthetase fusion protein (neuA1), and putative glycosyltransferase
(waaV) genes, complete cds; heptosyltransferase II (waaF) gene,
partial cds; and unknown genes.
AY044156
AY044156.1 GI:14993916
Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
Campylobacteraceae; Campylobacter.
1 (bases 1 to 13484)
Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E.,
Michniewicz,J., Cunningham,A.M. and Wakarchuk,W.W.
The Genetic Bases for the Variation in the Lipo-oligosaccharide of
the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF
SIALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE
J. Biol. Chem. 277 (1), 327-337 (2002)
11689567
2 (bases 1 to 13484)
Gilbert,M., Michniewicz,J., Karwaski,M.-F., Cunningham,A. and
Wakarchuk,W.W.
Direct Submission
Submitted (06-JUL-2001) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A
0R6, Canada
Location/Qualifiers
1..13484
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43429"
/serotype="0:1"
/db_xref="ATCC:43429"
/db_xref="taxon:197"
<1..348
/gene="waaC"
<1..348
/gene="waaC"
/codon_start=1
/transl_table=1
/product="heptosyltransferase I"
/protein_id="AAK73176.1"
/db_xref="GI:14993917"
/translation="FAKEVLNLSGIDETHIELAPKFNLELMFAFTKQMDLIIGNDSGP
THLAFALNKASITIFGATPSYRNAPQTHINKIIDTGKIKQNAKHIDKDFCITRIEE
DIFKLAAGLUNEX"
320..1225
/gene="htrB"
320..1225
/gene="htrB"
/codon_start=1
```



Percent Similarity:	69.3%	Conservative:	45
Best Local Similarity:	53.9%	Mismatches:	72
Query Match:	51.4%	Indels:	18
DB:	15	Gaps:	6
US-10-734-719-9 (1-291) x AY044156 (1-13484)			
Qy	5	IlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsnAsp	24
Db	7128	TTGGTATCGGTAAATGGACCTAGTTTAAAAAATATAGATTATAAAGGTACCAAAACA	7187
Qy	25	PheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyLeuGlyLys	44
Db	7188	TTTGATGTATTAGATGCAATCAGTTTATTTTGAAGATAGATATTTTGTGGGTAAAGAT	7247
Qy	45	CysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHis	64
Db	7248	GTAAATATATGTTTTTTTAAATCCTTTTGTGTTTTTGAACAATACTACTAGTAAAAA	7307
Qy	65	LeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAla	84
Db	7308	CTTATACAAAACGAAGAATATAACATAGAAAATATAGTTTGTCTACTATTAAATTAGAA	7367
Qy	85	HisLeuGlnAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeu	104
Db	7368	TACATAGATGATTCAATTTGTTGATAATTTTGAATTATATCTTTAGTGATCTTTTITA	7427
Qy	105	GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle	124
Db	7428	GGTCATGAGATTATAAAAAAGCTTAAAGATTTTGTGCTATATTAAATATAATGAAATA	7487
Qy	125	TyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly	144
Db	7488	TACAATAGGCAAGAAATTACTAGTGGAGTTATATGTGTGTACAGCTGTGGCTCTGGGA	7547
Qy	145	TyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer---TyrAla	163
Db	7548	TATAAAAGTATATATATAGTGGTATGATTTTATCAAGATACAAATAATTATATGCC	7607
Qy	164	PheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSer	183
Db	7608	TTCGATAATAATAAAAAAATTTATTAATAAATCAATGCACGTGTTTAAAAATCAAAATTC	7667
Qy	184	HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGlyLysThr	203
Db	7668	AAATTTATCAATCATCAATGCGATGCTGATTTTACAAGCTTTAGATTATCTAATGAAAGG	7727
Qy	204	TyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeu	223
Db	7728	TATGATGGAATATTATTATAGTTTA-----AATTC-----GATGAGTATTTTAAATG	7775
Qy	224	AlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys---AsnAsnTyrThrLysAsp	242
Db	7776	GCTCGGATATTGGAAGTGATTTTGTATTGACTAAAAAACCAAAATAATATAATGAT	7835
Qy	243	IleLeuIleProSerSerGluAla-----TyrGlyLysPheSerLysAsnIle	258
Db	7836	ATATTGATACCAGATAAGTATGTCACAGAAAGATATTATGGA-----	7877
Qy	259	AsnPheLysLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeu	278
Db	7878	-----AAAAAAGTAGGTTAAAAAATAATCCCAATTATAGCTAATAAAAGATCTAATT	7931
Qy	279	ArgLeuProSerAspIleLysHisTyrPheLysGlyLys	291
Db	7932	AGACTTCACAGTACATAAAACACTACTTAAGGGAANA	7970
RESULT 73			
CJ11168X4			
LOCUS			
DEFINITION			
segment 4/6.			
ACCESSION			
AL139077			

VERSION	AL139077.2	GI:6968444
KEYWORDS	Campylobacter jejuni subsp. jejuni NCTC 11168	
SOURCE	Campylobacter jejuni subsp. jejuni NCTC 11168	
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.	
REFERENCE	1 Parkhill, J., Wren, B.W., Mungall, K., Ketley, J.M., Churcher, C., Basham, D., Chillingworth, T., Davies, R.M., Feltwell, T., Holtroyd, S., Jagels, K., Karlyshev, A.V., Moule, S., Pallen, M.J., Penn, C.W., Quail, M.A., Rajandream, M.A., Rutherford, K.M., van Vliet, A.H., Whitehead, S. and Barrrell, B.G.	
TITLE	The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences	
JOURNAL	Nature 403 (6770), 665-668 (2000)	
PUBMED	10688204	
REFERENCE	2 (bases 1 to 282183)	
AUTHORS	Parkhill, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-FEB-2000) Submitted on behalf of the Campylobacter sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk	
COMMENT	Notes: Details of C. jejuni sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/C_jejuni/).	
FEATURES	Location/Qualifiers	
source	1..282183 /organism="Campylobacter jejuni subsp. jejuni NCTC 11168" /mol_type="genomic DNA" /strain="NCTC 11168" /db_xref="taxon:192222" complement(43..1443) /gene="Cj1009c" complement(43..1443) /note="Cj1009c, unknown, len: 466 aa; 29.1% identity to HP0282" /codon_start=1 /transl_table=1 /product="hypothetical protein Cj1009c" /protein_id="CAB73265.1" /db_xref="GI:6968445" /db_xref="GOA:Q9PNT1" /db_xref="InterPro:IPR006037" /db_xref="UniProtKB/TrEMBL:Q9PNT1" /translation="MNNLIIDGILAKHFLERLCFEKGLGYFTVVQCONSEKNLNI SSEYIDLHVFDPTSTARLEINSKDFKQAFIYMODEFTKKSVEALSNPLNLEIEM DFWGLSVNDTHANLADARMTLSRRFMDLFDIALTAQYIGLGVGEIMEVKIPAGSIFA YRHISIQOKRWIRVLVYRNSKIYFKVPSFVLEPNDILIVGDPVQLQSPHNIRGKA GQFPFPGSNVFALIDMKNNQMRVLDLTKLTKQSKNAKREFFIHVNPGLGVWYE KKLKLSDEKGEVFFDYFNTDFKQISMWLQNNDIGLVTDIKNFEKEKQAFDLKIPIM KYGEASFDELKEAIIILSADESELENNANVTIDLSKQLDFGLVLYYNPNSQNTDMEE YFRSLKLYDKNIOIKNKDENPLNLQVREDLLQFVFSQKELNLRDFARNLSNLNR HYKRNQYQLFIVE"	
gene	1473..2594 /gene="tgt"	
CDS	1473..2594 /gene="tgt" /RC_number="2.4.2.29" /note="Cj1010, tgt, probable queueine tRNA-ribosyltransferase, len: 373 aa; similar to many e.g. TGT_ECOLI queueine tRNA-ribosyltransferase (EC 2.4.2.29) (375 aa), fasta scores; opt: 1016 z-score: 1190.9 E): 0, 41.3% identity in 373 aa overlap. 56.7% identity to HP0281"	
gene	/codon_start=1 /transl_table=1	
CDS	/product="queueine tRNA-ribosyltransferase" /protein_id="CAB73266.1" /db_xref="GI:6968446" /db_xref="GOA:Q9PNT0" /db_xref="InterPro:IPR002616"	



```

/db_xref="InterPro:IPR004803"
/db_xref="UniProtKB/Swiss-Prot:Q9PNT0"
/translation="MEFKLKHGQMAKVCBETTIAHSTFLTPVMPVGTGCAVKSLDAN
DMKNELDAKILANTHYMLRPTSKVVDFFGHSTFKDPRSFLTDSGGFOAFSLSKN
SHFNIEGIEPKSHIDRSILFTPKSLDTPQDFNSDITMIMLIDDLVALPATKRVKISV
DRITLWAKEAITVHKMKNKGIGTQNI FGI IOGGTIDYERKRKCALSLNMPDPGLAI
GGISVEENALTEVTQNLNYPYDENPRVLMGVGTPEIDLVENVERGVDMFDCVMPTR
NARGPPTFSFGFNKKAEIFINDHEVIDSTCSYCTCRNEDSFYLNHLFKAKELTFPR
LASLHNLHYTLKARKREALNNSFTQFRNFYHLRGK"
2596.3363
/gene="Cj1011"
CDS
2596.3363
/gene="Cj1011"
/translation="Cj1011, probable membrane protein, len: 255 aa; no
Hp match"
/codon_start=1
/product="putative membrane protein"
/protein_id="CAB73267.1"
/db_xref="GI:6968447"
/db_xref="GOA:Q9PNS9"
/db_xref="InterPro:IPR002523"
/translation="MUELHNLLKILQAKNLETFYSBIYGOKIFVYVGLNLETLWFND
EKYKLDGDFKLSLIEFSNFIKSLIDEPKQVNTHPQNLLEKGIILKGGFKNFY
KKSFLVQKINKQLKQINLSEAFNLLLSQAQYKXKHLKILNLSISILSKNTEHLAR
IDLYLTNTAIKNEKMSIYLLSISSIFLPLNLIVGFGMTNNTNLFKDDSPYGTLY
IFSLICILVIFGYFYYSKTKTFDLDEGKAKKQTK"
3416.3492
/gene="trRNA-Arg"
tRNA
3416.3492
/gene="trRNA-Arg"
/product="trRNA-Arg"
/translation="trRNA Arg anticodon GCG, Cove score 83.29"
complement(3495.3944)
/gene="Cj1012c"
complement(3495.3944)
/gene="Cj1012c"
/translation="Cj1012c, possible membrane protein, len: 149 aa; no
Hp match"
/codon_start=1
/translation="Cj1012c, possible membrane protein, len: 149 aa; no
Hp match"
/product="putative membrane protein"
/protein_id="CAB73268.1"
/db_xref="GI:6968448"
/db_xref="UniProtKB/TrEMBL:Q9PNS8"
/translation="MVFCAALALLPLGVIAVLYITSDGKTKTKKQTSOKOHVAEKT
KFPDLDKMLIAASDVKLTDIELKELAKLYVQTHKLGSKTKELDEAKKLEFVSA
LAANINASQTVSYLNKELKKIYGYKKEIDAYEHMGLAKRKIKEDK"
complement(3961.7206)
/gene="Cj1013c"
complement(3961.7206)
/gene="Cj1013c"
/translation="Cj1013c, probable membrane protein, len: 1081 aa;
contains three domains; aa 1-90 contains three membrane
spanning domains; aa 90-780 is non-membrane, and aa
780-1081 contains ten possible membrane spanning domains.
Some similarity in C-terminal membrane domain to eukaryotic
proteins e.g. CCSA CHLRE cytochrome C biogenesis protein
CCSA (353 aa), fasta scores; opt: 459 z-score: 515.4 E():
2.1e-21, 36.0% identity in 267 aa overlap, and to
prokaryotic equivalents e.g. CCMP_ECOLI cytochrome C-type
biogenesis protein CCMP (647 aa); BLASTP scores; E = 0.20,
24% identity in 293 aa overlap. 39.0% identity to HP0378
(domains 2+3 only)"
/codon_start=1
/product="putative membrane protein"
/protein_id="CAB73269.1"
/db_xref="GI:6968449"
/db_xref="GOA:Q9PNS7"
/db_xref="InterPro:IPR002541"
/db_xref="UniProtKB/TrEMBL:Q9PNS7"
/translation="WKNIKISGDLRVSVLFLALFALCATATFIESAYGTPTAWMV
YDFWREYIQLLIGLNLGMRFKGLKPLMFIHISFLFVLGVSAMTRYAGPEG
ILPIRETKLISSSTSLRISAIKDGERYSAVNDRYIGNLPEANFKLNLGDDQ
AVLKQDILNLHATYYKENNNDPLVLMLQSKQSGQVDVREFKEGVNIEGVNFAFM
NDNVKAFKIDENLITLSSSENHLFLSMDGQNLDLKIGKANAKERRUYEINDLSFV
KKAASLHAQEALEGNRPQDSFWLWFKSAWLEVGRTWLIITFFGPQNNKNSLLHLFK
DQALSNNKLELTGSNALKLELSYKNESKEFYIPEYNKPMIELADQKFFTSWLSY
EOPFDLILRDLRVDPGMSPAYSEITVKNNNFDYRI FNNVLDYDGYRKYQ
SSVDDEKGTVLNKPQPKIPTVIGYELLCGLGEMFNFLPHSRFTLRLINKDTLK
HTSVIIFILLSPGSEKTPAQDLNSTLTPVNTNHAALATLIVQKADGRWVPFDTLS
REILEKHQSDSYGQNSAVMLSNLVDVDRWQLEPFLIMPQNAVRDAIANLEIPS
AKTAYKDFDENRNYKLYQVVENARNRPNARGVDFKEIITKLDNRANVNLVSGEL
FKFTPVQNNPNNWLAPSAVTTLKGDGCHIVLALIQNYFSAVENAFKQGNMIRADEG
LKFTKEYOEKIGYKVPMSKTVEMEIFSNKAEI FVKLAPVYIAGLLILVFSKMGV
PNLKISIPFKVYVNLVAFVHTVGLGLRAVLGSHAPMSNGESMVYTAWALSQVI
FFSRKSPALSLSILSGVLMVAHLSMNFOITNLVPLNSYWLISIHVSVTASVGF
LGLCALLIGFTFLMCFUKDKGKNLNLIRNITRATINEMAMI FGLCULTVGNFLGA
IWANESGRYMSWDSKETVALSVILVYAAIHLRMI PKYCNQVFALMSMFAYWVLIIM
TYFGVNYFLTGLHGYAAGEAAQIPNYVYWGFMALMVVLAF FARRKRNFKVL"
complement(7224.7919)
/gene="livF"
CDS
complement(7224.7919)
/gene="livF"
/translation="Cj1014c, livF, probable branched-chain amino-acid
ABC transport system ATP-binding protein, len: 231 aa;
similar to e.g. LivF_ECOLI high-affinity branched-chain
amino acid transport ATP-binding protein (237 aa), fasta
scores; opt: 686 z-score: 797.6 E(): 0, 46.4% identity in
233 aa overlap. No Hp ortholog. Contains PS00017
ATP/GTP-binding site motif A (P-loop), PS00211 ABC
transporters family signature, and Pfam match to entry
PF00005 ABC_tran"
/codon_start=1
/translation="Cj1014c, livF, probable branched-chain amino-acid
ABC transport system ATP-binding protein, len: 231 aa;
similar to e.g. LivF_ECOLI high-affinity branched-chain
amino acid transport ATP-binding protein (237 aa), fasta
scores; opt: 686 z-score: 797.6 E(): 0, 46.4% identity in
233 aa overlap. No Hp ortholog. Contains PS00017
ATP/GTP-binding site motif A (P-loop), PS00211 ABC
transporters family signature, and Pfam match to entry
PF00005 ABC_tran"
/product="branched-chain amino-acid ABC transport system
ATP-binding protein"
/protein_id="CAB73270.1"
/db_xref="GI:6968450"
/db_xref="GOA:Q9PNS6"
/db_xref="InterPro:IPR003439"
/db_xref="UniProtKB/TrEMBL:Q9PNS6"
/translation="WLI VKDLHVYVGLIEAVKGI DFKVETGHVLSIGNSAGKTKSTL
NALLNSVKRTGEVNFKLDYTKRHLTVLQKGI ALVPSRRVFINLSVENLKI GAFN
NAENYHREQNYKFLPRLAKHALAGTSGGEAQMALSAISALMSEPKLLMLDEPSL
GLAPKIVGEVFDIIVRLKEGITTILLVEQNASALKISDYAVVLENGRIVMENEAKNL
IGDDEIRKKYLGL"
complement(7299.7841)
/gene="livF"
misc_feature
/translation="Cj1014c, livF, probable branched-chain amino-acid
ABC transport system ATP-binding protein, len: 231 aa;
similar to e.g. LivF_ECOLI high-affinity branched-chain
amino acid transport ATP-binding protein (237 aa), fasta
scores; opt: 686 z-score: 797.6 E(): 0, 46.4% identity in
233 aa overlap. No Hp ortholog. Contains PS00017
ATP/GTP-binding site motif A (P-loop), PS00211 ABC
transporters family signature, and Pfam match to entry
PF00005 ABC_tran"
/product="branched-chain amino-acid ABC transport system
ATP-binding protein"
/protein_id="CAB73270.1"
/db_xref="GI:6968450"
/db_xref="GOA:Q9PNS6"
/db_xref="InterPro:IPR003439"
/db_xref="UniProtKB/TrEMBL:Q9PNS6"
/translation="WLI VKDLHVYVGLIEAVKGI DFKVETGHVLSIGNSAGKTKSTL
NALLNSVKRTGEVNFKLDYTKRHLTVLQKGI ALVPSRRVFINLSVENLKI GAFN
NAENYHREQNYKFLPRLAKHALAGTSGGEAQMALSAISALMSEPKLLMLDEPSL
GLAPKIVGEVFDIIVRLKEGITTILLVEQNASALKISDYAVVLENGRIVMENEAKNL
IGDDEIRKKYLGL"
complement(7299.7841)
/gene="livF"
misc_feature
/translation="Cj1014c, livF, probable branched-chain amino-acid
ABC transport system ATP-binding protein, len: 231 aa;
similar to e.g. LivF_ECOLI high-affinity branched-chain
amino acid transport ATP-binding protein (237 aa), fasta
scores; opt: 686 z-score: 797.6 E(): 0, 46.4% identity in
233 aa overlap. No Hp ortholog. Contains PS00017
ATP/GTP-binding site motif A (P-loop), PS00211 ABC
transporters family signature, and Pfam match to entry
PF00005 ABC_tran"
/product="branched-chain amino-acid ABC transport system
ATP-binding protein"
/protein_id="CAB73270.1"
/db_xref="GI:6968450"
/db_xref="GOA:Q9PNS6"
/db_xref="InterPro:IPR003439"
/db_xref="UniProtKB/TrEMBL:Q9PNS6"
/translation="WLI VKDLHVYVGLIEAVKGI DFKVETGHVLSIGNSAGKTKSTL
NALLNSVKRTGEVNFKLDYTKRHLTVLQKGI ALVPSRRVFINLSVENLKI GAFN
NAENYHREQNYKFLPRLAKHALAGTSGGEAQMALSAISALMSEPKLLMLDEPSL
GLAPKIVGEVFDIIVRLKEGITTILLVEQNASALKISDYAVVLENGRIVMENEAKNL
IGDDEIRKKYLGL"
complement(7299.7841)
/gene="livF"
misc_feature

```

## Alignment Scores:

Pred. No.: 2.16e-90 Length: 282183  
Score: 804.00 Matches: 158  
Percent Similarity: 69.3% Conservative: 45  
Best Local Similarity: 53.9% Mismatches: 72  
Query Match: 15.4% Indels: 18  
DB: 15 Gaps: 6

US-10-734-719-9 (1-291) x CJ111168X4 (1-282183)

Qy 5 llelleAlaGlyAsnGlyProSerleuLysGluIleAspTyrSerArgLeuProAsnApp 24

Db 134487 TTGGTATCGGTAATGGACCTAGTTTAAAAAATATAGATTATAAAGGGTTTACCAAAACAA 134546

Qy 25 PheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLys 44

Db 134547 TTTGATGATTATTAGATGCAATCAGTTTATTATTGAGATAGATATTTTGGTGAAGAT 134606

Qy 45 CysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeuLysHis 64

Db 134607 GTAAATATGTTTATTTTAACTCTTTTGTGTTTTTGAACATACTACTACTAGTAAAAA 134666

```
Qy 65 LeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAla 84
Db 134667 CTTATACAAACGAAGAATAAATAGTTTGTCTACTATTAAATTAGAA 134726
Qy 85 HisLeuGluAsnGluAsnPhelValysThrPheTyrAspTyrPheProAspAlaHisLeu 104
Db 134727 TACATAGATGATTTCAATTTGTTGATAATTTTGAATTATATCTTTTAGTGATCGCTTTTAA 134786
Qy 105 GlyTyrAspPhePheLeuGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIle 124
Db 134787 GGTCAATGAGATTATAAAAGCTTAAGATTTTTTGGCTATATTAAATATTAATGAAATA 134846
Qy 125 TyrPheAsnGlnAtgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGly 144
Db 134847 TACAATAGCGCAAGAATTTACTAGTGGAGTTTATATGTGTGTACAGCTGTGGCTCTGGGA 134906
Qy 145 TyrIysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSer---TyrAla 163
Db 134907 TATAAAGTATATATATAAGTGGTATTGATTTTATCAAGATACAAATAAATTTATATGCC 134966
Qy 164 PheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSer 183
Db 134967 TTCGATATAATAAAAAAATTTATTAATAAATCAATGCGTGTGTTTAAAAATCAAAATTC 135026
Qy 184 HisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr 203
Db 135027 AAATTTATCAATCATTCATGCGCATGCTGATTTTACAAGCTTTAGATTAATCTAATGAAGG 135086
Qy 204 TyrIysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhelIleGluLeu 223
Db 135087 TATGATGGAATATTTATAGTTTA-----AATTCA-----GATGAGTATTTTAAATTG 135134
Qy 224 AlaProAsnLeuAsnSerAsnPhelIleLeuGlnGluLys---AsnAsnTyrThrLysAsp 242
Db 135135 GCTCGGATATTGGAAGTGATTTTGTATTGAGTAAAAAACCACCAAAAAATATATAAATGAT 135194
Qy 243 IleLeuIleProSerSerGluAla-----TyrGlyLysPheSerLysAsnIle 258
Db 135195 ATATTGATACCAGATAAGTATGCAACAAAGATATTATGGA----- 135236
Qy 259 AsnPhelLysLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeu 278
Db 135237 -----AAAAAAGTAGGTTAAAGAAAAATCTCCATTATAGCTTAATAAAGATCTAATT 135290
Qy 279 ArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
Db 135291 AGACTTCAAGTGACATAAAACACTACTTTAAAGGAAAAA 135329
```

## RESULT 74

CP000057\_09/c

## WPCOMMENT

Sequence split into 20 fragments LOCUS CP000057 Accession CP000057

Fragment Name	Begin	End
CP000057_00	1	110000
CP000057_01	100001	210000
CP000057_02	200001	310000
CP000057_03	300001	410000
CP000057_04	400001	510000
CP000057_05	500001	610000
CP000057_06	600001	710000
CP000057_07	700001	810000
CP000057_08	800001	910000
CP000057_09	900001	1010000
CP000057_10	1000001	1110000
CP000057_11	1100001	1210000
CP000057_12	1200001	1310000
CP000057_13	1300001	1410000
CP000057_14	1400001	1510000
CP000057_15	1500001	1610000
CP000057_16	1600001	1710000
CP000057_17	1700001	1810000
CP000057_18	1800001	1910000
CP000057_19	1900001	1913428

Continuation (10 of 20) of CP000057 from base 900001 (CP000057 Haemophilus influenzae)

Alignment Scores:

Pred. No.:	3 57e-83	Length:	110000
Score:	744.50	Matches:	149
Percent Similarity:	67.7%	Conservative:	50
Best Local Similarity:	50.7%	Mismatches:	86
Query Match:	47.6%	Indels:	9
DB:	15	Gaps:	6

US-10-734-719-9 (1-291) x CP000057\_09 (1-110000)

```
Qy 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluLeuAspTyrSerArgLeu 21
Db 81735 AAGTCGTGTCATTATTGTCAGGTAAATGCAACAAAGTTTAAATCAATGACTATAGTTTATTA 81676
Qy 22 ProAsnAspPheAspValPheAlaCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41
Db 81675 CCTAAAGATTATGATGTTTTTCGGTTGCAATCAATTTTATTTTGAGGATCATTTATTTCTT 81616
Qy 42 GlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr 61
Db 81615 GGTAAAGAAATAAAAGGTATTTTTTAATGTTCTGTAAATTTTGAACAATACTATACG 81556
Qy 62 LeuLysHisLeuIleGlnAsnGlnGluTyrGlu---ThrGluLeuIleMetCysSerAsn 80
Db 81555 TTTATGCAATTAATTTAAATAATGAATATGAATATGCTGATATTTATTTATCATCTTTT 81496
Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 81495 CTAAATTTAGGGGATTTCAGAAATTTAAAGAAA---ATCCAGCGTTTAAAGAAAATTTACTACCA 81439
Qy 101 AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPhelys 120
Db 81438 CAAATCGATCTTGCTGATAGCTATTTAAAAAACTACGAGCTTTGATGCTCATTTACAA 81379
Qy 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 81378 TATCAGCAACTATATGAGAAATAGAGAGATTACATCAGGCGCTATATGTGTGCAAGTGCA 81319
Qy 141 IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGly 159
Db 81318 ACTGCAATGGTTTATAAAGATCTTTATTTGACAGGCATTCGATTTTATCAAGAAAAAGGG 81259
Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPhelys 179
Db 81258 AATCCTTAGCGATTTTCATCATCAACACAGAAAAATATTATTAATTTATTTACCTTCTTTTCA 81199
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 81198 CAAATATAAAGTCAAAAGCGATATCCATTTATGGAATATGATTTTAAATGCACTTTATTTT 81139
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 81138 TTACAAAAACATTTATGGAGTAAATATTATTGCAATTCGCAGAAAGTCTCTCTATGTAAT 81079
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPhelIleGlnGluLysAsn 237
Db 81078 TATTTCCCTTTATCACCA---CTGAATAACCCCAATTTACTTTTATTTCTCGAAGAAAAAGAA 81022
Qy 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257
Db 81021 AATTACACACAAGATATTTTAAATTCGCGCAAGATTGTTGATATAAAAAAATTTGGT----- 80968
Qy 258 IleAsnPhelLysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277
Db 80967 ---ATATATCCAAACCAAGAATTTACCAAAATCTGATTTTTCGGTTGATCGGATATA 80911
Qy 278 LeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291
Db 80910 TTACGTTTACCTTAATGATATAAAACACGCTTTTGAAGCAAAAG 80869
```

RESULT 75

CQ873070/c  
 LOCUS CQ873070 191996 bp DNA linear PAT 27-SEP-2004  
 DEFINITION Sequence 683 from Patent WO2004078949.  
 ACCESSION CQ873070  
 VERSION CQ873070.1 GI:52746846  
 KEYWORDS  
 SOURCE Haemophilus influenzae  
 ORGANISM Haemophilus influenzae  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.  
 REFERENCE 1  
 AUTHORS Bakaletz, L.O., Munson, R.S. and Dyer, D.W.  
 TITLE Genes of an otitis media isolate of nontypeable haemophilus influenzae  
 JOURNAL Patent: WO 2004078949-A 683 16-SEP-2004;  
 CHILDREN'S HOSPITAL, INC. (US)  
 FEATURES  
 source 1. 191996  
 /organism="Haemophilus influenzae"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:727"  
 ORIGIN  
 Alignment Scores: Length: 191996  
 Pred. No.: 6,11e-83 Matches: 149  
 Score: 744.50  
 Percent Similarity: 67.7% Conservative: 50  
 Best Local Similarity: 50.7% Mismatches: 86  
 Query Match: 47.6% Indels: 9  
 DB: 2 Gaps: 6  
 US-10-734-719-9 (1-291) x CQ873070 (1-191996)  
 QY 2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21  
 DB 43859 AAGTGTCTATTATGAGTAATGGAACAAAGTTTAAATCAATGACTAGTTATTA 43800  
 QY 22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrIleu 41  
 DB 43799 CCTAAGATTATGATGTTTCCGTGCAATCAATTTATTTTGGAGGATCAATTTCTT 43740  
 QY 42 GlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrThr 61  
 DB 43739 GGTAAAGAAATAAAGGATTTTAAATGTTCTGTAATTTTGAACAATCACTATACG 43680  
 QY 62 LeuLysHisLeuIleGlnAsnGlnLysGlu---ThrGluLeuIleMetCysSerAsn 80  
 DB 43679 TTTATGCAATTAATAAATAATGAATGCTGATATTTATTTATCATCTTTT 43620  
 QY 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100  
 DB 43619 CTAAATTTAGGGATTTCAGAAATTAAGAAA---ATCCAGCGCTTTAGAAAAATTAACCA 43563  
 QY 101 AsnAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120  
 DB 43562 CAATCGATCTTGCTAGCTATTTTAAATAAATACAGCGCTTTGATGCTCATTTACAA 43503  
 QY 121 PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140  
 DB 43502 TATCAGCACTATATGAGATAAAGGATTAACATGCGCTCTATATGTGTCAGTGCA 43443  
 QY 141 IleAlaLeuGlyTyrLysGluIleTyrIleuSerGlyIleAspPheTyrGln---AsnGly 159  
 DB 43442 ACTGCAATGGTTATAAAGATCTTTATTTGACAGCATTTGATTTTATCAAGAAAAAGGG 43383  
 QY 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLys 179  
 DB 43382 AATCCTTACCGATTTTCATCATCAACACAGAAATATTTAAATATTATTAACCTTCTTTCA 43323  
 QY 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199  
 DB 43322 CAAATATAAGTCNAAGCGATATCCATCTCTATGGAATATGATTTAAATGCACTTTATTTT 43263

QY 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219  
 DB 43262 TTACAAACATTATGAGTAATAATTTATTCATTTCCGAGAAAGTCTCTATGTAAT 43203  
 QY 220 PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsn 237  
 DB 43202 TATTTCCCTTTATCACCA---CTGAATAACCAATTTACTTTTATTTCTCGAAGAAAAAGAA 43146  
 QY 238 AsnTyrThrLysAspIleLeuProSerSerGluAlaTyrClyLysPheSerLysAsn 257  
 DB 43145 AATTACACACAGATATTTTAAATTCGCGCAAGTTTGTATATAAAAAAATTGGT----- 43092  
 QY 258 IleAsnPheLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277  
 DB 43091 ---ATATATTCCAAACCAAGATTTTACCAAAATCTGATTTTCGGTTGATCTGGATATA 43035  
 QY 278 LeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291  
 DB 43034 TTACGTTTACCTATATGATATAAACACACGCTTTGAAAGCAAG 42993  
 RESULT 76  
 LOCUS BD134499  
 DEFINITION Lipopolysaccharide alpha-2,3 sialyltransferase of Campylobacter jejuni and its uses.  
 ACCESSION BD134499  
 VERSION BD134499.1 GI:23229444  
 KEYWORDS JP 2002507424-A/1.  
 SOURCE Campylobacter jejuni  
 ORGANISM Campylobacter jejuni  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.  
 REFERENCE 1 (bases 1 to 1293)  
 AUTHORS Gilbert, M. and Wakarchuk, W.W.  
 TITLE Lipopolysaccharide alpha-2,3 sialyltransferase of Campylobacter jejuni and its uses  
 JOURNAL Patent: JP 2002507424-A 1 12-MAR-2002;  
 NATIONAL RESEARCH COUNCIL OF CANADA  
 COMMENT OS Campylobacter jejuni  
 PD JP 2002507424-A/1  
 PD 12-MAR-2002  
 PF 22-MAR-1999 JP 2000538012  
 PR 20-MAR-1998 US 60/078891,18-MAR-1999 US 09/272960 PI  
 MICHEL GILBERT, WARREN W WAKARCHUK  
 PC  
 C12N15/09, C12N1/21, C12N5/10, C12N9/10, C12P19/26, C12N15/00, C12N5/ PC  
 00  
 CC Campylobacter jejuni OH4384 cst-I gene alpha2,3- CC  
 sialyltransferase  
 FH Key Location/Qualifiers  
 FT CDS  
 FEATURES  
 source 1. 1293  
 /organism="Campylobacter jejuni"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:197"  
 ORIGIN  
 Alignment Scores: Length: 1293  
 Pred. No.: 7,6e-85 Matches: 1293  
 Score: 743.00 Conservative: 147  
 Percent Similarity: 66.0% Mismatches: 45  
 Best Local Similarity: 50.5% Indels: 81  
 Query Match: 47.5% Gaps: 5  
 DB: 2  
 US-10-734-719-9 (1-291) x BD134499 (1-1293)  
 QY 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
 DB 46 ATGCAAAATATAATCATAGCAGAAATGACCTAGCTTAAATAATTAATTAATAAAGA 105  
 QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40

Db	106	CTGCCTAGAGAATATGATGTTTATAGGTGTACCAAGTTTATTTTGAAGATAAGTATTAT	165
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	166	TTAGGAAAAAGATTAAAGCAGTATTTTAAATCTCGTGTCTTTTACAAAGATATCAC	225
Qy	61	ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn	80
Db	226	ACTGCAAAACACTTATATACTAAAAATGAGTATGAATAAAAAATATTTTGGCTCTACA	285
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	286	TTTAATTTACCTTTTATTGAAAGCAATGATTTTATACATCAATTTTATAATTTTTCGCC	345
Qy	101	AspAlaHisLeuGlyTyrAspPheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	346	GATGCAAACTTGGCTATGAAATTTTAAAGCACTTTAAAGCACTTTTATGCTATATAAAA	405
Qy	121	PheHisGluLeuGluLeuAlaProAsnLeuAsnSerAsnPheIleGlnGluLys---	236
Db	694	GCAAAATCATTTTCCCTTTATCAATTAATAATAACAATTTTACATTTAGAAAAAAGCAT	753
Qy	237	AsnAsnTyrThrLysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLys	256
Db	754	ATAAATCTATAAATGATATTTTATTGACTGATAATACTCTCGCGTAAGTTTTTATAAAA	813
Qy	257	Asn-----LysileLys	266
Db	814	AATCAACTTAAAGCTGATAATAAAATATGCTTTAATTTTATAATAATTTCTTCATAA	873
Qy	267	GluAsnValTyrTyrLysLeuLeuLysAspLeu	277
Db	874	GATAATTAAATTTTAAACAAAGAAATTT	906
RESULT 77	AR473907	LOCUS	
DEFINITION	Sequence 1 from patent US 6689604.	DNA	linear PAT 20-FEB-2004
ACCESSION	AR473907		
VERSION	AR473907.1	GI:42712446	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1293)		
AUTHORS	Gilbert, M. and Wakarchuk, W. W.		
TITLE	Lipopolysaccharide alpha.-2,3 sialyltransferase of Campylobacter jejuni and its uses		
JOURNAL	Patent: US 6689604-A 1 10-FEB-2004; National Research Council of Canada; Ottawa; CAX;		
FEATURES	Location/Qualifiers		

source	1..1293	/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Alignment Scores:			
Pred. No.:	7.6e-85	Length:	1293
Score:	743.00	Matches:	147
Percent Similarity:	66.0%	Conservative:	45
Best Local Similarity:	50.5%	Mismatches:	81
Query Match:	47.5%	Indels:	18
DB:	2	Gaps:	5
US-10-734-719-9 (1-291) x AR473907 (1-1293)			
Qy	1	MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg	20
Db	46	ATGCAAAATATATATCATAGCAGAAATGGACCTAGCCTAAAAAATATATATATAAAGA	105
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	106	CTGCCCTAGAGATATGATGTTTTTAGGTGTAAACAGTTTTATTTTGAAGATAAGTATTAT	165
Qy	41	LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr	60
Db	166	TTAGGAAAAAGATTAAAGCAGTATTTTAAATCTCGTGTCTTTTACAACAGTATCAC	225
Qy	61	ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn	80
Db	226	ACTGCAAAACAACTTATATACTAAAAAATGAGTATGAATAAAAAATATTTTGTCTACA	285
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100
Db	286	TTTAATTTACCTTTTATTGAAAGCAATGATTTTATACATCAATTTTATAATTTTCCCC	345
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	346	GATGCAAACTTGGCTATGAAATTTTAAAGCACTTTAAAGCACTTTTATGCTATATAAAA	405
Qy	121	PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	406	TACAATGAAATTTATTTCAATAAAGAAATTTACTTCGGCGCTCTATATGTGTGCAATTGCT	465
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	466	ATTGCAATAGGATATATAAACCACTATTTATGTGGCAATGATTTTATGAAGGAGATGTT	525
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaPro-----Asp	177
Db	526	ATTATCTCTTTTGAAGCTATGATACAAATATATAAAACAATCTTCTCGGAATAAAGAT	585
Qy	178	PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu	197
Db	586	TTCAAACTTCAAAATGTT-----CATTTAAGGAATACGATATAGAAGCAATTA	633
Qy	198	GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu	217
Db	634	AAATGTTTAAATCAATATACAAAGTTAATATCTACGCATTTGTGTGATGATTTCTATTG	693
Qy	218	AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys---	236
Db	694	GCAAAATCATTTTCCCTTTATCAATTAATAATAACAATTTTACATTTAGAAAAAAGCAT	753
Qy	237	AsnAsnTyrThrLysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLys	256
Db	754	ATAAATCTATAAATGATATTTTATTGACTGATAATACTCTCGCGTAAGTTTTTATAAAA	813
Qy	257	Asn-----LysileLys	266
Db	814	AATCAACTTAAAGCTGATAATAAAATATGCTTTAATTTTATAATAATTTCTTCATAA	873
Qy	267	GluAsnValTyrTyrLysLeuLeuLysAspLeu	277

```
Db      874  GATAAATTAAATTAATTTTAAACAAAGAAATT 906
RESULT 79
AR489130
LOCUS   AR489130
DEFINITION Sequence 1 from patent US 6709834.
ACCESSION AR489130
VERSION   AR489130.1
KEYWORDS  AR489130.1 GI:47256032
SOURCE   Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1293)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Lipopolysaccharide .alpha.-2,3 sialyltransferase of campylobacter jejuni and its uses
JOURNAL Patent: US 6709834-A 1 23-MAR-2004;
National Research Council of Canada; Ottawa;
CAX;
FEATURES             Location/Qualifiers
     source          1..1293
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
Alignment Scores: 7.6e-85 Length: 1293
Pred. No.: 743.00 Matches: 147
Score: 66.0% Conservative: 45
Percent Similarity: 50.5% Mismatches: 81
Best Local Similarity: 47.5% Indels: 18
Query Match: 2 Gaps: 5
DB:

US-10-734-719-9 (1-291) x AR489130 (1-1293)
Qy      1  MetLysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
Db      46  ATGCAAAATATATATAGCAGAGAAATGGACCTAGCCTAAATAATATTAATTAAGA 105
Qy      21  LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr 40
Db      106  CTGCTAGAGAAATATGATGTTTATAGGTGTAAACAGTTTATTTTGAAGATAAGTATAT 165
Qy      41  LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db      166  TTAGGAAAAAGATTAAAGCAGTATTTTAAATCCTGCTGTCTTTTACACACAGTATCAC 225
Qy      61  ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80
Db      226  ACTGCAAAACAACCTATATCTAAATAATGAGTATGAAATAATAATAATTTTGTCTTACA 285
Qy      81  TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db      286  TTTAATTACCTTTTATGAAGCAATGATTTTACATCAATTTTATAATTTTCCCC 345
Qy      101  AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db      346  GATGCAAAACTTGCTATGAAGTATTTCAAACCTTAAAGAAATTTATGCTTATATAAAA 405
Qy      121  PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db      406  TACAATGAAATTTATTTCAATAAAAGAAATTAATCTCGGCGCTATATATGTGCAATGCT 465
Qy      141  IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db      466  ATTGCATTAGATATATAAACCATCTATTATGTGGCATTGATTTTATGAAGGAGATGTT 525
Qy      161  SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaPro-----Asp 177
Db      526  ATTATCTCTTTGAAGCTATGAGTACAAATATAAAAAACAATCTTCTCCTGGAATAAAGAT 585
Qy      178  PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu 197
Db      586  TTCAAACCTTCAAATTTGT-----CATTTAAGGAATACGATATAGAAGCATTA 633
Qy      198  GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu 217
Db      634  AAAATGTTTAAATCAATATACAAAGTTAATATCTACGCATTTGTGATGATCTTATTTG 693
Qy      218  AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys-- 236
Db      694  GCAATCATCTTTCCTTTATCAATTAATATATAACAATTTACATTAGAAAATAAGCAT 753
Qy      237  AsnAsnTyrThrLysAspIleLeuProSerSerGluAlaTyrGlyLysPheSerLys 256
Db      754  AATAATCTATAAATGATATTATTAATGACTGATAATACTCTCGCGCTAAAGTTTATAAAA 813
Qy      257  Asn-----11164 bp DNA linear BCT 27-OCT-2004
Db      814  AATCAACTTAAAGCTGATAATAAATATGCTTAATTTTATAATATCTTCAATCTTAA 873
Qy      267  GluAsnValTyrTyrLysLeuIleLysAspLeu 277
Db      874  GATAATTTAATTAATTTTAAACAAAGAAATT 906
RESULT 79
AF130466
LOCUS   AF130466
DEFINITION Campylobacter jejuni OH4384 peptide chain release factor 2 (prfB) gene, partial cds; alpha-2,3-sialyltransferase (cst-I), sulfate adenyltransferase subunit CysD (cysD), sulfate adenyltransferase subunit CysN (cysN), putative sodium/sulfate symporter (sodium sulfate symporter), putative adenylsulfate kinase (cysC), hypothetical protein, putative glycosyltransferase, and hypothetical protein genes, complete cds; and cjl457c gene, partial cds.
ACCESSION AF130466
VERSION   AF130466.2
KEYWORDS  GI:54639977
SOURCE    Campylobacter jejuni
ORGANISM  Campylobacter jejuni
REFERENCE 1 (bases 1 to 1164)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Lipopolysaccharide .alpha.-2,3 sialyltransferase of campylobacter
JOURNAL Patent: PCT WO99/49051 18-MAR-1999;
REFERENCE 2 (bases 1 to 1164)
AUTHORS Gilbert,M., Brisson,J.R., Karwaski,M.F., Michniewicz,J., Cunnigham,A.M., Wu,Y., Young,N.M. and Wakarchuk,W.W.
TITLE Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
REFERENCE Identification of the glycosyltransferase genes, enzymatic synthesis of model compounds, and characterization of nanomole amounts by 600-mhz (1)h and (13)c NMR analysis
JOURNAL J. Biol. Chem. 275 (6), 3896-3906 (2000)
PUBMED 10660542
REFERENCE 3 (bases 1 to 1164)
AUTHORS Gilbert,M., Michniewicz,J. and Wakarchuk,W.W.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1999) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada
REFERENCE 4 (bases 1 to 1164)
AUTHORS Gilbert,M., Brochu,D. and Karwaski,M.-F.
TITLE Direct Submission
JOURNAL Submitted (25-OCT-2004) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A 0R6, Canada
REMARK Sequence update by submitter
COMMENT On Oct 26, 2004 this sequence version replaced gi:6469928.
FEATURES             Location/Qualifiers
     source          1..1164
                     /organism="Campylobacter jejuni"
                     /mol_type="genomic DNA"
                     /strain="OH4384"
                     /db_xref="taxon:197"
```

```

gene
<1..812
/feature="Penner serotype 19"
/feature="prfb"
CDS
<1..812
/feature="prfb"
/feature="Helicobacter pylori Prfb homolog"
/codon_start=3
/transl_table=11
/product="neptide chain release factor 2"
/protein_id="AAF13494.1"
/db_xref="GI:6469929"
/translation="ALFNDAPKLEDTITSLISMSSGNDGKNAIVSIHPGAGCTES
NDASILRYMLRFRCEBGFVETLDFQEGEAGLKDVSLVKGENAYGYLKAENGIIH
RLVTSRPSDAGRHTSSVVMVSPELDDETEIEEKDIRIDYVRASAGGQHNKT
ESAVRITHPTGIIVQCONDRSQHNKATAFKMLKSLRYELEMLMKQDSANTGKSEI
GWHGQIRSYLVFPYQQVQDNRSGEAFSQVDNILDGDIKMKIEGVLIALKAE"
812..2104
/feature="cst-I"
812..2104
/feature="cst-I"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAF13495.1"
/db_xref="GI:6469930"
/translation="WTRTRMENELIVSKMQLNIAGNPSLKNINYKELPREYDVER
SNDFYEDQYILGKKIKAVFPNPGVFLQOYHTAKOLILKNEYEIKNIFCSTFNLPIE
CNDLHQYFNFPDAKLYEVLENLKEFYAYIKNEYIYFNKRITSQVVMCAIALGY
KTIYLCGIDFVEGDYFPEAMSTNIKTIPFGIKDFKPSNCHSKEYDEALKLALSIY
KNTIYALCDDGILANHFPLSININNFTLENKHNNSINDILLTNTDTPGVSVFKNLQKA
DNKIMLNFYNTLHSDNLIKELNKEIAVLKQTTORAKARQNHLISYKLGQALILNSK
YVLGFLSLPFIILSIVISHKQEQKAYKFKVKNPNLALPLETYPDYNEALKEKCEFT
YKLGEEFTKAGKNVYGEYIKFIFKDVPRLKRPEKGE"
2105..3010
/feature="cysD"
2105..3010
/feature="cysD"
/feature="similar to Escherichia coli CysD"
/codon_start=1
/transl_table=11
/product="sulfate adenylyltransferase subunit CysD"
/protein_id="AAF13496.1"
/db_xref="GI:6469931"
/translation="MIFNLTHLRQLENSIFIMREVISEFEPKPMLYSVGKDSSVNLH
LQKAFYTPPLPLVHVDTTWKEMIEFRDKRAKELGMELIVYQNPKIKELNLSPF
THDSSMHTDISKTGLQMLDLYQDFVGGARRDEEKRAKERIYSFRDENHTWDPK
NORPELWLYNRRHKKGSIRVPLSNWTELDIWOYIYKENIPIPSLYFAKRSVIEY
MGKILVDDDRMPKELAKNSKEEMVRFITLGCYPLTGAINSNASNVLIEIBELLSKT
SERQRLDITDEEASMEKKKEGYF"
3010..4428
/feature="cysN"
3010..4428
/feature="cysN"
/feature="similar to Escherichia coli CysN"
/codon_start=1
/transl_table=11
/product="sulfate adenylyltransferase subunit CysN"
/protein_id="AAF13497.2"
/db_xref="GI:54639978"
/translation="MQTNIEYLOPHENKELCRFTTCGVDDGKSTLIGLLYDTPKAL
FSDQLSTLEKOSKMGAGDKLDLFDLVDGLASEREQGITIDVATRFITSNKRKFIIA
DTPGHEQYTRNMTAGSTADAILIDARKGLVKQTKRHSYIVSLGKFNFIANMK
DLVSYEEKIFNNIKDYKIPYLOEDIQTHFIPTICALNGENITOKSRNLSWGKTEL
LALLDEIKINKIIONDFIMPVQYNRHLNFRSCFNIAAGVYKLODRIIVLPSPQKS
KIKSITINDKDLTDLNENILPSONEACVNMVSLCLEDEIDISRGDILASINHLK
MSNAFEMIIWMSQQLDNENYLIKRAHNLCNVKFNINYYKDIINTPEENASNLIL
NDIANCTLUKDLNKLKEYKNTGLSGFIIDKYSNETNLAAGMIILNLSQQSQRIYT
QAEIELNAFIRKNYPWECRKI"
4425..6137
/feature="sodium sulfate symporter"
4425..6137
/feature="sodium sulfate symporter"
/codon_start=1

gene
6145..6744
/feature="cysC"
6145..6744
/feature="cysC"
/codon_start=1
/transl_table=11
/product="putative adenylyltransferase kinase"
/protein_id="AAF36523.1"
/db_xref="GI:54639980"
/translation="MDKMLTWHNTITKEORSKLKQKPCVLMTLGLSASGKSTIANA
LEQKLHEGFHTYLLDGDNIIRGLNKLDFGDNRSVENIRRIAEELCKLFDVDSGLIILS
AATSPKSDRALARSIVKKDEFIEIFVDTPLECEKRDPKGLYKARNGIKNFTGID
SYEPKQNPPIHLQNSKAKFNENVDIILNFKENGIDYI"
6836..8011
/feature="variable poly(G) tract"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF36524.1"
/db_xref="GI:54639981"
/translation="MPKDYHDKYGCIFIHVPKVAGTSIERVVFETDKWLGVHVRALDY
INQDNKTESYFSAFVRNPFDRMVSAYFLYKGGNGNGDKIWADENKNDFTEQOYV
LALKNNIKDKILSWQHTPQYKFCIDENKNILVNETIGLENINNDPKIKVQNELNFOR
NIHNSSSKHEIFSYNENKTYNIIAKLYKEDTDFDYDLEYKESYKMLDVQFLNLM
YKEKLFSKNKEIEKRLSQFKNKINSQNNIILQOTNQIHNLTLENKNQLLITKE
NLLFNQNNYKAKTRVQNSLYKQALILNSKSVGLFSPLFIILSIISHKQEQKA
YKFKVKNPNLALPLETYPDYNEALKEKCEFTYKLGEEFTKAGKNWYGGGIIILAL"
7390*7991
/feature="results in in extended C-terminus hypothetical
protein"
/replaces="g"
8048..9778
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAF36525.1"
/db_xref="GI:54639982"
/translation="MKTGVGVPIYNNVEKYLRCLDSVNNQTYKNLQVLVNDGSTDE
NSLNTAKETLADERILFDKENGILSSRNVGIEFFSKYEDFKNIQOELKENSILVER
KLDNEDNPNIYKIKSSNPFKNKDELLNFKAPDIDYILFDSDDYWELNCEBCEVPR
MGQEVVWFDFYTFYDIENPPKQIKTILEDYQFKSETITSQWLEKTIENNFATFW
LQCMCMINPIQFNHILKFIINGIHHEDHFGMLLCLQANKIYINLKNKLYIRVRPNS
IMNYNDNGKNINSKLNFLNLDVIGKYYKILSYGINAFALNFSNNFNKDLI
KLFNAFKNCEENWIYDIAQYPTNDRSLFIEIFRMKNYETNENLILDFIAMIIN
NNKITIVKOSNEIQNNNTIKLYCEKINSQNNIILQOTNQIHNLTLENKNQLLITK
ENLNFQNNYKAKTRVQNSLYKQALILNSKSVGLFSPLFIILSIISHKQEQK
AYKFKVKNPNLALPLETYPDYNEALKEKCEFTYKLGEEFTKAGKNWYGGGIIKFWL
IDIQNLKRNK"
9778..10737
/feature="variable poly(G) tract; similar to N-termini of
hypothetical protein Cj1433c from C. jejuni NCTC 11168"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF36526.1"
/db_xref="GI:54639983"
/translation="MKLQWLLSLVKKLOIKKKYKHILEQSRITELKINILNNM
TTGGGITEKIKKITTFRPFLHGISFHLAEHCNLCNCFSCDNFSAQKAGYYELNPEKDV

```



```
/transl table=11
/product="CysD"
/protein_id="AAK33820.1"
/db_xref="GI:60679076"
/translation="MFNMLTHQLNENISFIMREVISPEFEPKAMLYSGVKDSSVNLH
LLQKAFYTPVPLVHVDTTWKFKEMIEFRDKRAKEIGMELIVYONPKIKELNLSPP
THDSMTDIKSTQGLKQMLDYOQDAVFGGARDEEKSRAKERIYSFRDENHTWDPK
NORPELWNLNGRHKHKGSIIRVPLSNWTELDIMQYIKENIPIPSLYFAKRSVIEY
MGAKILVDDDDMPKELAKNSKEEMVRFRTLCYPTJTGAINSNASNVLEIIBELLKST
SERQRLIDTDEEASMEKKKEGYF"
2323..3741
/gene="cysN"
2323..3741
/gene="cysN"
/notes="putative sulfatase subunit"
```

gene

```
/transl table=11
/product="CysN"
/protein_id="AAK33821.1"
/db_xref="GI:60679077"
/translation="MQTNIEKYLQEHENKELCRFITCGSVDDGKSTLIGRLLYDTKAL
FDSQSTLEKQKMGNGDGLDFALLVDGLASEREQGITIDVATRFITSNKRKFIIA
PTDQEQTRMTACSTADIAIILIDARKGVLTQTKRHSYIVSLGKFNFIATNKM
DLVSEYEEKFNIIQNDYKPIYLOEDIQTHFIPICALNGENITQKSNLSYKGETL
LALDEIKINKIIQNDFTMPQVYVNRPHLFRSCFNAGSYVLODBIIVLPWMOKS
KIKSTINDIKDLSTLDENEIIPSONEACVMAVSLCEDEIDISRGDILASINHDLK
MSNAFAMIWMSQOLDNENYLIKRAHNLNVKFNINYYKKOINTKEENASNLIL
NDIANCTLKDKNALKEYKDNKTLGSLFIIDKYSNETLAAGMIKIILNSQOSRIYT
QAEIELNAFIRKNTPEWGRKI"
3738..5450
/codon_start=1
/transl table=11
/product="putative sodium/sulfate symporter"
```

CDS

```
/protein_id="AAK33822.1"
/db_xref="GI:60679078"
/translation="WKIIVGITLILLILRNKIKPAVLFGSLAGFYALGYLDPKT
WISYTWDSLISLMLLIVASTVEKTIIEWASKFIQKNYNLSLRIGVITCAVSAP
LNTAVASFMGIVKNNKQAPSKLLIPLSYPAIVGGVITLVTSTNLIINSFVQNG
LPSLUMDFYFYGIALSGIMLVLMIFSLKPHYENKDKNIKEHLIALKLVNLSLIG
KSIENKRLNIEFLSVIEIORNQSITPVSHNEIINAKDTLIFSGDISQLETLNFKDG
KLADGVELKSFIDAIISPTSLNIGSKVEANFRSKFDAIISLQRGDVHKKIGE
VLQAGDRMILAVGDFTRDLNLAKNFYLLSKONEKLDTKKSLITTMAFLSVTVFS
ALNCISTKALLICLGLLELLCKFIKLDIEKRFPFDIIFIIYGVSSLIATTKVLAVDSGLAK
DLAQITTFPGHYGIVGIFVIGLITLTLTEFITNNAALAFPIGFAIQAALGVNPL
PFIIFAYAGSAGFMIPHYQTHLMVSSLCDYKMTDFIKIGAIIVSIVYSAIVLIGVPL
IFNF"
5458..6057
/gene="cysC"
5458..6057
/gene="cysC"
/codon_start=1
/transl table=11
/product="putative adenylyl sulfatase kinase"
```

gene

```
/protein_id="AAK33823.1"
/db_xref="GI:60679079"
/translation="MDKNLTWHNTNITKEQSKLKNQKPCVLMVLTGLSASGKSTIANA
LEQKHLGPHFTYLLDGNIRNMLNKGIDGFDSNRVENIRIAELCKLFDVDSGLIILS
AFISPFKSDRALARSLVKKDEFIIFVDTPLELCEKPKDKPKGLKKARNGEIKNFTGID
SPYKPNPETHLQNSKAKFNENVDIILNLYLKENGIDYT"
6149..7351
/notes="non-experimental evidence, no additional details recorded"
```

CDS

```
/note="contains variable poly(G) tract"
/codon_start=1
/transl table=11
/product="hypothetical protein"
/protein_id="AAK33824.1"
/db_xref="GI:60679080"
/translation="MFKDYHDKYGCIFIHVPKVAGSTIERVVVFETDKVLGVHVRALDY
INODKNKPFVSFYSFVRNPFDRMVSAFHYLKKGGNGNDKIWADENCKLFDPTPEQFV
LALKNKIKOKILSKOQHTPOYKFCIDENKILNLYNFKLENINNDKFIKVNELNFDK
NLHNSKSHIEIFSNYNEKTYNIIAKLYKEDFTLFDYDLEYKESIYKNLDVQFLNLM
YKEFLSKNKEIEKLRLSQFKKNKEINSQNNIIIQQTNQIHNLTNTLENKNKQLITKE
```

```
NLLNFQNNYKAKTRVQNOLSYKLGQALILNSKSVLGFSLPFIILSIISHKQEQKA
YKFKVKNPNLALPPLETYPDYNEALKEKECFYKLGALIQASKNMYGGGDSYCLV
EFLSSIKN"
7304..7305
/misc_difference
/notes="results in extended C-terminus hypothetical protein"
/replaces="g"
7304
/misc_difference
/notes="results in truncated hypothetical protein"
/replaces=""
```

CDS

```
7362..9092
/codon_start=1
/transl table=11
/product="putative glycosyltransferase"
/protein_id="AAK33825.1"
/db_xref="GI:60679081"
/translation="MKTGVVPIYVNEKYLRCLDSVVNQTYKNLOVLVNDGSTDE
NSLNIATKEYTLKDERFLDKNGSLSSNRNVGIEFSEYDFKNITOEKENSLSVEF
KLDNEDNPYNIYKIKSNFKNKDELLKAPDIDYIIFLSDSDYWELNCEIECVPR
MDGVEVWFDDYFYDDIENPKQIETILEDYQFKSETITSQWLEKTLNENFTAFW
LQOMCINFIQPLNHLKAFINGI IHEDHFGMLLCLQANKIYINLNKLYIYVRPNS
INWYNDGNKINKSLNFCNLNLNVIDGKKYKILSYGINAFALNFSNFHNDLLI
KLFNAKFNCEENIYDIIAQPTNDLSLFIIFRIMKNYETNENLIDFTIAMIN
NNKTIIVKQSNIEQNNNTKIYCEKINSQNNIILOQTNOIHNLTNTLENKQLLITK
ENLLNFQNNYKAKTRVQNOLSYKLGQALILNSKSVLGFSLPFIILSIISHKQEQK
AYKFKVKNPNLALPLETYPDYNEALKEKECFYKLGALIQASKNMYGGGYYIKFWL
IDIQNLKRN"
```

CDS

```
9092..9241
/inference="non-experimental evidence, no additional details recorded"
/notes="contains variable poly(G) tract"
/codon_start=1
/transl table=11
/product="hypothetical protein"
/protein_id="AAK33826.1"
/db_xref="GI:60679082"
/translation="MKLKQWLLSLVKKLQIKKKYKHILEQSRITELEKINNLLNNM
TTGGY"
```

misc\_difference

```
9236..9237
/notes="results in extended C-terminus hypothetical protein"
/complement(10043..>10484)
/locus_tag="Cj1457c"
/complement(10043..>10484)
/locus_tag="Cj1457c"
/codon_start=2
/transl table=11
/product="Cj1457c"
/protein_id="AAK33827.1"
/db_xref="GI:60679083"
/translation="KRMIQIYIKISKEAEKELKKQEQFFKLLKGEVLGHYFGPKCFLOE
DLSAELERPRARDISAMGLLIGVAYETGEGLALNLENEIFKDALEKPAKQVSRREM
WRYLEELKWRDYDEKAHFCIEFFLQKGSYATVVLLEILHTIFS"
```

gene

CDS

## ORIGIN

## Alignment Scores:

Pred. No.:	1,2e-83	Length:	10484
Score:	740.50	Matches:	148
Percent Similarity:	66.4%	Conservative:	40
Best Local Similarity:	52.3%	Mismatches:	86
Query Match:	47.4%	Indels:	9
DB:	15	Gaps:	4

US-10-734-719-9 (1-291) x AY791515 (1-10484)

Qy	1	MetLysLysValleilleAlaGlyAsnGlyProSerIeuLysGluIleAspTyrSerArg	20
Db	170	ATGCAAAATATATCATACAGAGAAATGGACCTAGGCTAAAAAATATATATATAAAGA	229
Qy	21	LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyr	40
Db	230	CTGCCTAGAGAAATATGATGTTTTTAGGTGTACACAGTTTTTATTGAGATAGTATTAT	289



```

Qy 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyr 60
Db 290 TTAGGAAAAAGAAATTAAGCAGTATTTTTTAATCTCGTGTCTTTTCAACAGATATCAC 349

Qy 61 ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn 80
Db 350 ACTGCAAAACAACCTTATCTATAAAATGAGTATGAATAAAATAATTTTTTGCTCTACA 409

Qy 81 TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro 100
Db 410 TTTAAATTTACCTTTTATTGAAGCAATGATTTTTTATCAATTTTATTAATTTTTTCCCC 469

Qy 101 AsnAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys 120
Db 470 GATGCAAAACCTTGCTATGAAGTATTGAAACCTTTAAAGAAATTTTATGCTTATATAAA 529

Qy 121 PheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla 140
Db 530 TACAATGAATTTATTTTCATATAAAGAAATTTACTTCGGCGCTCTATATGTGCAATGCT 589

Qy 141 IleAlaLeuGlyTyrLysGluLeuTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer 160
Db 590 ATTGCATTTAGCATATAAACCATCTATTATGTGGCATTGATTTTATGAAGGAGATGTT 649

Qy 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaPro-----Asp 177
Db 650 ATTATATCTTTTGAAGCTATGAGTACAAATATAAAACAATCTTCTCGAATAAAGAT 709

Qy 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu 197
Db 710 TTCAAACCTTCAATGCT-----CATCTAAGGAATACGATATAGAGCATTA 757

Qy 198 GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu 217
Db 758 AAATTTGTTAAATCAATATACAAAGTTAAATATCTACGCATTTGTGTGATGATCTTATT 817

Qy 218 AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLys--- 236
Db 818 GCAAAATCATTTTCTCTTATCAATTAATTAATAACAATTTCACTTAGAAAATAAGCAT 877

Qy 237 AsnAsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLys 256
Db 878 AATAATTCATAAATGATATTTTATGACTGATATACTCTCTGGCGTAAGTTTATATAAA 937

Qy 257 AsnIleAsnPheLysLysIleLysLysGluAsnValTyrTyrLysLeu---IleLys 275
Db 938 AATCAACTTAAAGCTGATATAAAATTAATGCTTAATTTTATATAATATCTTCATTCTAAA 997

Qy 276 AspLeuLeu 278
Db 998 GATACTTTA 1006

RESULT 81
CS222474
LOCUS CS222474 960 bp DNA linear PAT 15-DEC-2005
DEFINITION Sequence 4489 from Patent WO2005111066.
ACCESSION CS222474
VERSION CS222474.1 GI:83683477
KEYWORDS Haemophilus influenzae
SOURCE Haemophilus influenzae
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Haemophilus.
REFERENCE 1.
AUTHORS Masignani V.
TITLE Polypeptides from non-typeable haemophilus influenzae
JOURNAL Patent: WO 2005111066-A 4489 24-NOV-2005;
Chiron SRL. (IT)
FEATURES
source
1. 960
/organism="Haemophilus influenzae"
/mol_type="unassigned DNA"

```

```

/db_xref="taxon:727"
/note="Name: NTH2503"

ORIGIN
Alignment Scores:
Pred. No.: 2,26e-83 Length: 960
Score: 730.50 Matches: 150
Percent Similarity: 64.8% Conservative: 47
Best Local Similarity: 49.3% Mismatches: 80
Query Match: 46.7% Indels: 27
DB: 2 Gaps: 7

US-10-734-719-9 (1-291) x CS222474 (1-960)

```

```

Qy 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
Db 79 AAGTCTGTCATTTATGTCAGGTAATGGAACAAGTTTAAATCAATGACTATAGTTATTATA 138

Qy 22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41
Db 139 CCTAAGATTTATGATGTTTCCGTTGCAATCAATTTTATTTTGAAGATCATTTATTTCT 198

Qy 42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr 61
Db 199 GGCAAGAAAAATAAAAAAGGTATTTTAAATGCTCTGTAATTTTGAACAATACTATACG 258

Qy 62 LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu-----LeuIleMetCys 78
Db 259 TTTATGCAATTAATTAATAAATGAATATGAATATGAATATGCTGATATATCTTAGCA 318

Qy 79 SerAsnTyrAsnGlnAla-----HisLeuGluAsnGluAsn 90
Db 319 TCTTTTCTGAATTTAGGGGATTCACATTAAGAAAAATCCAGCATTTAGAAAAA----- 372

Qy 91 PheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys 110
Db 373 -----TTACTGCCACAAATTGATCTTTGGTCATTGCTATTGAAA 411

Qy 111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130
Db 412 AAATCTACGAGCTTTTAAATGCTCATTTACAATATCAGAAATTAATGAGAAATAAGAGATT 471

Qy 131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150
Db 472 ACATCAGCGTTTATATGTGTCAGTGGCGACTGCAATGGGTTATAAGATCTTTATTATTA 531

Qy 151 SerGlyIleAspPheTyrGln---AsnGlySerSerTyrAlaPheAspThrLysGlnGlu 169
Db 532 ACAGTATTGATTTTATCAAGAAAAAGGAATCCTTACGCATTTTCATCATCAAAAAAGAA 591

Qy 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSer 189
Db 592 AATATATTAAATTTATACCTTCTTTTTCACAAAAATAAAGTCAAAAGCATATCCATTCT 651

Qy 190 LysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyr 209
Db 652 ATGGAATATGATTTAAATGCACTTTATTTTACAAAAACATTTATGGAGTAAATATTAT 711

Qy 210 CysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSer 229
Db 712 TGCATTTCCGCGAGAAAGTCTCTGTATTTATTTTCTTTTATCACA---CTGAATAAC 768

Qy 230 -----AsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuProSer 247
Db 769 CCAATTTGCTTTTATCCAGAGAAAAAATAATACACAAGATATTTTAATTCGCGCG 828

Qy 248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysGlu 267
Db 829 AAGTTTGATATAAAAAAATTTGGT-----ATATATTCCAAACCAAGATTTTACCAA 879

Qy 268 AsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyr 287
Db 880 AATCTGATTTTTCGGTTGTTCTGGGATATATTTACGTTTACCTTAATGATATAAACACGCC 939

```

```

QY      288 PheLysGlyLys 291
      |||
      |||
      |||
Db      940 TTAATAATCAAGA 951

RESULT 82
CQ872953/c
LOCUS      13379 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 566 from Patent WO2004078949.
ACCESSION CQ872953
VERSION    CQ872953.1 GI:52746760
KEYWORDS
SOURCE
ORGANISM   Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE
AUTHORS    Bakletz, L.O., Munson, R.S. and Dyer, D.W.
TITLE       Genes of an otitis media isolate of nontypeable haemophilus
            influenzae
JOURNAL     Patent: WO 2004078949-A 566 16-SEP-2004;
            CHILDREN'S HOSPITAL, INC. (US)
FEATURES
source     1..13379
            /organism="Haemophilus influenzae"
            /mol_type="unassigned DNA"
            /db_xref="taxon:727"

ORIGIN
Alignment Scores:
Pred. No.:      2.88e-82      Length:      13379
Score:          730.50      Matches:      150
Percent Similarity: 64.8%      Conservative: 47
Best Local Similarity: 49.3%      Mismatches: 80
Query Match:    46.7%      Indels:      27
DB:            2      Gaps:      7

US-10-734-719-9 (1-291) x CQ872953 (1-13379)

QY      2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
      |||
      |||
      |||
Db      11607 AAGTCGTGCAATTTGCGAGTAAATGGAACAAGTTTAAATCAATGACTATAGTTTATTA 11548

QY      22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41
      |||
      |||
      |||
Db      11547 CCTAAGATTATGATGTTTCCGTTGCAATCAATTTATTTTGAAGATCAATTAATTTCTT 11488

QY      42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr 61
      |||
      |||
      |||
Db      11487 GCGAAGAAAAATAAAAGGTAATTTTAAATGTTCTGTAATTTTGAACAATACATATACG 11428

QY      62 LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu-----LeuIleMetCys 78
      |||
      |||
      |||
Db      11427 TTTATGCAATTAATTAATAAATGAATATAATGAATATGCTGATATATCTTAGCA 11368

QY      79 SerAsnTyrAsnGlnAla-----HisLeuGluAsnGluAsn 90
      |||
      |||
      |||
Db      11367 TCTTTTCTGAATTTAGGGGATTCAACATTAAGAAAAATCCAGCATTTAGAAAAA----- 11314

QY      91 PheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys 110
      |||
      |||
      |||
Db      11313 -----TTACTGCCACAAATATGATCTGGTCATCTGGTCATCTATTGAAA 11275

QY      111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130
      |||
      |||
      |||
Db      11274 AAACACAGCGTTTAAATGCTCAATTAACAATATACGAATATATACGAATAAGAGGATT 11215

QY      131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150
      |||
      |||
      |||
Db      11214 ACATCAGCGGTTTATGTGTGCGAGTGCAGTGCATGCAATGGGTTATAAAGATCTTTATTTA 11155

QY      151 SerGlyIleAspPheTyrGln---AsnGlySerSerTyrAlaPheAspThrLysGlnGlu 169
      |||
      |||
      |||

```

```

Db      11154 ACAGGTATTGATTTTATCAAGAAAAAGGAATCCTTACCATTTTCATCATCAAAAAAGAA 11095
QY      170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSer 189
      |||
      |||
      |||
Db      11094 AATATTATTAAATATTACCTTCTTTTTCACAAAAATAAAAGTCAACGCGATATCATCTCT 11035

QY      190 LysAsnThrAspIleLysAlaLeuGluLysThrTyrLysIleLysLeuTyr 209
      |||
      |||
      |||
Db      11034 ATGGAATATGATTTAAATGCACTTTATTTTACAAAAACATTATGAGTAATAATTTAT 10975

QY      210 CysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSer 229
      |||
      |||
      |||
Db      10974 TGCATTTCCGACAAAAGTCTCTATGTAATTTATTTCTTTTATCACCA---CTGAATAAC 10918

QY      230 -----AsnPheIleIleGlnGluLysAsnAspTyrThrLysAspIleLeuIleProSer 247
      |||
      |||
      |||
Db      10917 CCAATTGCTTTTATTCGAGAAAAAGAAAAATACACACAGATATTTTAATTCGCGCG 10858

QY      248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysGlu 267
      |||
      |||
      |||
Db      10857 AAGTTTGTATATAAAAAAATTGGT-----ATATATTCCAACCAAGATTTTACCAA 10807

QY      268 AsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyr 287
      |||
      |||
      |||
Db      10806 AATCTGATTTTTCGGTTGTTCTGGGATATATTACGTTTACCTAATGATATAAAACAGCC 10747

QY      288 PheLysGlyLys 291
      |||
      |||
      |||
Db      10746 TTAATAATCAAGA 10735

RESULT 83
CQ000057_04/c
WFCOMMENT
Sequence split into 20 fragments LOCUS CQ000057 Accession CQ000057
Fragment Name      Begin      End
CP000057_00      1      110000
CP000057_01      100001 210000
CP000057_02      200001 310000
CP000057_03      300001 410000
CP000057_04      400001 510000
CP000057_05      500001 610000
CP000057_06      600001 710000
CP000057_07      700001 810000
CP000057_08      800001 910000
CP000057_09      900001 1010000
CP000057_10      1000001 1110000
CP000057_11      1100001 1210000
CP000057_12      1200001 1310000
CP000057_13      1300001 1410000
CP000057_14      1400001 1510000
CP000057_15      1500001 1610000
CP000057_16      1600001 1710000
CP000057_17      1700001 1810000
CP000057_18      1800001 1910000
CP000057_19      1900001 1913428
Continuation (5 of 20) of CQ000057 from base 400001 (CQ000057 Haemophilus influenzae)

Alignment Scores:
Pred. No.:      2.21e-81      Length:      110000
Score:          730.50      Matches:      150
Percent Similarity: 64.8%      Conservative: 47
Best Local Similarity: 49.3%      Mismatches: 80
Query Match:    46.7%      Indels:      27
DB:            15      Gaps:      7

US-10-734-719-9 (1-291) x CQ000057_04 (1-110000)

QY      2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
      |||
      |||
      |||
Db      47456 AAGTCGTGCTATTGTCAGGTAATGGAACAAGTTTAAATCAATGACTATAGTTTATTA 47397

QY      22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41
      |||
      |||
      |||

```

Db	47396	CCTAAGATTATGATGTTTCCGTTGCAATCAATTTTATTTTGAAGATCATTTATTTCTT	47337	JOURNAL	Patent: WO 2004078949-A 772 16-SEP-2004;
Qy	42	GlyLysCysAlaValPheTyrThrProAsnPhePheGluGlnTyrThr	61	FEATURES	CHILDREN'S HOSPITAL, INC. (US)
Db	47336	GGCAAGAAATAAAAAAGGATTTTAAATGTTCTGTAATTTTGAACAATACTATACG	47277	source	1. 117576
Qy	62	LeuLysHisLeuLeuGlnAsnGlnGluTyrGlu	78		/organism="Haemophilus influenzae"
Db	47276	TTTATGCAATTAATTAATAATAATGAATATAAATATGAATATGCTGATATATCTTAGCA	47217		/mol_type="unassigned DNA"
Qy	79	SerAsnTyrAsnGlnAla	90	misc_feature	/db_xref="taxon:727"
Db	47216	TCCTTTTCGATTTAGGGATTCAACATTAAGAAATCCAGCATTTTAGAAAA	47163		/note="Old seq 684 to long. replaced by new seq 684, from 000.001 to 349.980"
Qy	91	PheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys	110	misc_feature	new seq 772, from 300.001 to 417.556"
Db	47162	TTACTGCCCAAAATTCATCTTGCTCATCTGCTATTTGAAA	47124		93603
Qy	111	GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgile	130		/note="n is a, c, g, or t"
Db	47123	AAACTAGAGCTTTAATGCTCTATTACAAATATCACGAATATATGAGATAAGAGATT	47064		93620
Qy	131	ThrSerGlyValTyrMetCysAlaValAlaLeuGlyTyrLysGluLeuTyrLeu	150		/note="n is a, c, g, or t"
Db	47063	ACATCAGCGCTTATATGTGTGAGTGGCTGCAATGGGTTATAAAGATCTTTATTA	47004		93622..93625
Qy	151	SerGlyLeuAspPheTyrGln---AsnGlySerSerTyrAlaPheAspThrLysGlnGlu	169	ORIGIN	
Db	47003	ACAGGTATTGATTTTATCAAGAAAAAGGAATCCTTACGCATTTTCATCATCAAAAAAGAA	46944	Alignment Scores:	
Qy	170	AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSer	189	Pred. No.:	2.35e-81
Db	46943	AAATTTATTAATTTATACCTCTCTTTTCACAAATAAAGTCAACGATATCCATCT	46884	Score:	730.50
Qy	190	LysAsnThrAspLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyr	209	Percent Similarity:	64.8%
Db	46883	ATGGAATATGATTTAATGACCTTTATTTTACAAAAACATTTATGGAGTAAATTTAT	46824	Best Local Similarity:	49.3%
Qy	210	CysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSer	229	Query Match:	46.7%
Db	46823	TGCATTTCCGAGAAAGTCTCTATGTAATTTATTTCTTTATCACCAC---CTGAATAAC	46767	DB:	2
Qy	230	-----AsnPhelIleGlnGluLysAsnTyrThrLysAspIleLeuLeuPheProSer	247	US-10-734-719-9 (1-291) x CQ873159 (1-117576)	
Db	46766	CCAATTCCTTTATTCAGAGAAAGAAATAATACACAAAGATATTTAAATTCGCCCG	46707	Qy	2
Qy	248	SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysIleLysGlu	267	Db	18786
Db	46706	AGTTTGTATATAAAAAATGGT-----ATATATCCAAACCAAGATTTACCAA	46656	Qy	22
Qy	268	AsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyr	287	Db	18786
Db	46655	AACTGATTTTTCGGTTGTTCTGGGATATATTACGTTTACCTAATGATATAAAACACGCC	46596	Qy	42
Qy	288	PhelLysGlyLys	291	Db	18666
Db	46595	TTAAATCAAGA	46584	Qy	62
RESULT 84				Db	18606
CQ873159/c				Qy	79
LOCUS	CQ873159	117576 bp	DNA	Db	18546
DEFINITION	Sequence 772 from Patent WO2004078949.		linear	Qy	91
ACCESSION	CQ873159			Db	18492
VERSION	CQ873159.1	GI:52746892		Qy	111
KEYWORDS				Db	18453
SOURCE	Haemophilus influenzae			Qy	131
ORGANISM	Haemophilus influenzae			Db	18393
	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			Qy	151
	Pasteurellaceae; Haemophilus.			Db	18333
REFERENCE				Qy	170
AUTHORS	Bakaletz, L.O., Munsen, R.S. and Dyer, D.W.			Db	18273
TITLE	Genes of an otitis media isolate of nontypeable haemophilus influenzae			Qy	190
				Db	18213
				Qy	210

```
Db 18153 TGCATTTCCGAGAAAGTCCTCTATGTAAATTATTTCTTTTATCACCA---CTGAATAAC 18097
Qy 230 -----AsnPhelIleGlnGluYsAsnAsnTyrThrLysAspIleLeuProSer 247
Db 18096 CCAATTGCTTTTATTTCCAGAGAAAGAAATATACACACAGATATTTTAAATCCGCG 18037
Qy 248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysLysIleLysGlu 267
Db 18036 AAGTTTGTATATAAAAAAATTGGT-----ATATATTCCAAACCAAGATTTACCAA 17986
Qy 268 AsnValTyrTyrLysLysAspLeuLeuArgLeuProSerAspIleLysHisTyr 287
Db 17985 AATCTGATTTTCGGTGTCTCGGATATATACGTTTACCTAAATGATATAAAACACGCC 17926
Qy 288 PheLysGlyLys 291
Db 17925 TTAAATAATCAAGA 17914

RESULT 85
CQ873071/c CQ873071 349980 bp DNA linear PAT 27-SEP-2004
LOCUS Sequence 684 from Patent WO2004078949.
DEFINITION CQ873071
ACCESSION CQ873071
VERSION CQ873071.1 GI:52746847
KEYWORDS
SOURCE
ORGANISM Haemophilus influenzae
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Haemophilus.
REFERENCE 1
AUTHORS Bakaletz,L.O., Munson,R.S. and Dyer,D.W.
TITLE Genes of an otitis media isolate of nontypeable haemophilus
influenzae
JOURNAL Patent: WO 2004078949-A 684 16-SEP-2004;
CHILDREN'S HOSPITAL, INC. (US)
FEATURES
source
1. 349980
/organism="Haemophilus influenzae"
/mol_type="unassigned DNA"
/db_xref="taxon:727"
misc_feature 93603
/note="n is a, c, g, or t"
misc_feature 93620
/note="n is a, c, g, or t"
misc_feature 93622..93625
/note="n is a, c, g, or t"
misc_feature 138581
/note="n is a, c, g, or t"

ORIGIN

Alignment Scores:
Pred. No.: 6 75e-81 Length: 349980
Score: 730.50 Matches: 150
Percent Similarity: 64.8% Conservative: 47
Best Local Similarity: 49.3% Mismatches: 80
Query Match: 46.7% Indels: 27
DB: 2 Gaps: 7

US-10-734-719-9 (1-291) x CQ873071 (1-349980)

Qy 2 LysLysValIleIleAlaGlyAsnGlyProSerLysLysGluIleAspTyrSerArgLeu 21
Db 318786 AAGTCTGTCATTTATTCAGGTAATGGAACAAGTTTTTAAATCAATGACTATAGTTTATTA 318727
Qy 22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu 41
Db 318726 CCTAAGATTTATGATGTTTTCGGTTGCAATCAATTTTATTTTGAAGATCAATTTTCTT 318667
Qy 42 GlyLysLysCysLysAlaValPheTyrThrProAsnPhePhePheGluGlnTyrTyrThr 61
Db 318666 GCAGAGAAATAAAAAAGGTATTTTTTAAATGTTCTGTAAATTTTGAACAATACTATACG 318607
```

```
Qy 62 LeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu-----LeuIleMetCys 78
Db 318606 TTTATGCAATTAATAAAAAATGAATATAAATAATGAATATGCTGATATATCTTAGCA 318547
Qy 79 SerAsnTyrAsnGlnAla-----HisLeuGluAsnGluAsn 90
Db 318546 TCTTTTCTGAATTTAGGGGATTCAACATTAAAGAAATCCAGCATTTAGAAAAA----- 318493
Qy 91 PheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys 110
Db 318492 -----TTACTGCCCAAAATTGATCTTGTCATCTGCTATTGAA 318454
Qy 111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIle 130
Db 318453 AAATACGAGCCTTTTAAATGCTCATTTACAATATCAGCAATATATATGAGATAAAGAGATT 318394
Qy 131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150
Db 318393 ACATCAGCGCTTTATATGTCAGTGGCAGCTGCAATGGGTTATATAAGATCTTTATTTA 318334
Qy 151 SerGlyIleAspPheTyrGln---AsnGlySerSerTyrAlaPheAspThrLysGlnGlu 169
Db 318333 ACAGGTATTTGATTTTATCAAGAAAAAGGAATCCTTTACGCATTTCAATCATCAAAAAGAA 318274
Qy 170 AsnLeuLysLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSer 189
Db 318273 AATATTTATTAATTTATTAATCTCTTTTTCACAAAATAAAGTCAAGCGATATCTCT 318214
Qy 190 LysAsnThrAspIleLysAlaLeuGluPheLysGluLysThrTyrLysIleLysLeuTyr 209
Db 318213 ATGGAATATGATTTAAATGCACCTTTATTTTACAAAAACATTATGAGTAAATATTAT 318154
Qy 210 CysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSer 229
Db 318153 TGCATTTCCGCAGAAAGTCTCTATGTAATTTATTTCTTTTATCACCA---CTGAATAAC 318097
Qy 230 -----AsnPhelIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuProSer 247
Db 318096 CCAATTGCTTTTATTTCCAGAGAAAGAAATAAATACACACAGATATTTTAAATCCGCG 318037
Qy 248 SerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysLysIleLysGlu 267
Db 318036 AAGTTTGTATATAAAAAAATTGGT-----ATATATTTCCAAACCAAGATTTACCAA 317986
Qy 268 AsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyr 287
Db 317985 AATCTGATTTTTCGGTGTCTCGGATATATTTACGTTTACCTAATGATATAAAACACGCC 317926
Qy 288 PheLysGlyLys 291
Db 317925 TTAAATAATCAAGA 317914

RESULT 86
BD426631_03/c
WPCOMMENT
Sequence split into 19 fragments LOCUS BD426631 Accession BD426631
Fragment Name Begin End
BD426631_00 1 110000
BD426631_01 100001 210000
BD426631_02 200001 310000
BD426631_03 300001 410000
BD426631_04 400001 510000
BD426631_05 500001 610000
BD426631_06 600001 710000
BD426631_07 700001 810000
BD426631_08 800001 910000
BD426631_09 900001 1010000
BD426631_10 1000001 1110000
BD426631_11 1100001 1210000
BD426631_12 1200001 1310000
BD426631_13 1300001 1410000
BD426631_14 1400001 1510000
BD426631_15 1500001 1610000
```

BD426631_16 1600001 1710000									
BD426631_17 1700001 1810000									
BD426631_18 1800001 1830121									
Continuation (4 of 19) of BD426631 from base 300001 (BD426631 The Nucleotide Sequence of									
Alignment Scores:									
Pred. No.:	3.13e-80	Length:	110000						
Score:	721.50	Matches:	145						
Percent Similarity:	66.7%	Conservative:	51						
Best Local Similarity:	49.3%	Mismatches:	89						
Query Match:	46.2%	Indels:	9						
DB:	2	Gaps:	6						
US-10-734-719-9 (1-291) x BD426631_03 (1-110000)									
Qy	2	LysLysVallelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu	21						
Db	79512	AAGTCGTCAATTATTCGAGGTAATGGACAAGTTTAAATCAATGACTATAGTTTATTA	79453						
Qy	22	ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrLeu	41						
Db	79452	CCTAAGATTATGATGTTTCCGTTGCAATCAATTTATTTTGAAGATCATATTATTTCTT	79393						
Qy	42	GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr	61						
Db	79392	GGCAAGAAATAAAAAAGATTTTTTAATGTTCTACAAATTTTGGACAAATCTATACT	79333						
Qy	62	LeuLysHisLeulleGlnAsnGlnGluTyrGlu--ThrGluLeulleMetCysSerAsn	80						
Db	79332	TTTATGCAATTAATTAATAATGAATATGAATATGCTGATATTAATTTATCATCTTTT	79273						
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100						
Db	79272	GTCAATTTAGGAGATTCAGAATTAAGAAA--ATTAAAAATGTACAAAATTTACTAACA	79216						
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120						
Db	79215	CAAGTTGATATGGACATTAATTAATTAACCAAGCTACCCGCTTTGATGCTCTATTTCACAA	79156						
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140						
Db	79155	TATAACGAATTATATGAATAAAGAGAATTACATCAGCGGTTTATATGTGTGAGTGCGCA	79096						
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln--AsnGly	159						
Db	79095	ACTGTAATGGTGTATAAGATCTTTATTAACAGGTATGATTTTATCAAGAAAAAGGG	79036						
Qy	160	SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys	179						
Db	79035	AATCCTTACGCAITTCATCATCAAAAAGAAATATTAATAAATTTATACCTTCTCTTTTCA	78976						
Qy	180	AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe	199						
Db	78975	CAAAATAAAGTCAAAGCGATATCCATCTATGGAATATGATTTAAATGCACTTTATTTT	78916						
Qy	200	LeuGluLysThrTyrLysIleLysLeuTyrCysLeuLysCysProAsnSerLeuLeuAlaAsn	219						
Db	78915	TTACAAAAAATATGAGGTAATAATTTATTTGCAATTCGCCAGAAAGTCCTCTATGTAAT	78856						
Qy	220	PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsn	237						
Db	78855	TATTTCTCTTATCACCA--CTGAATAACCAATTAATTTATTTCTCGAAGAAAGAAA	78799						
Qy	238	AsnTyrThrLysAspIleLeulleProSerSerGluAlaTyrGlyLysPheSerLysAsn	257						
Db	78798	AATTCACACAGATATTTTAATTTCCGCGAGTTTGTATATAAAAAATTTGGT-----	78745						
Qy	258	IleAsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeulleLysAspLeu	277						
Db	78744	--ATATATTCCAAAACCAAGAAATTCACAAATCTGATTTTTCGGTTCGATCTGGGATATA	78688						
Qy	278	LeuArgLeuProSerAspIleLysHisTyrPheLysGlyLys	291						

Db 78687 TTACGTTTACCTTAATGATATATAAAACACGCTTAAAAATCAAGA 78646

RESULT 87

AR274513\_03/c

WPCOMMENT

Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1830121

Continuation (4 of 19) of AR274513 from base 300001 (AR274513 Sequence 1 from patent U

Alignment Scores:				
Pred. No.:	3.13e-80	Length:	110000	
Score:	721.50	Matches:	145	
Percent Similarity:	66.7%	Conservative:	51	
Best Local Similarity:	49.3%	Mismatches:	89	
Query Match:	46.2%	Indels:	9	
DB:	2	Gaps:	6	

US-10-734-719-9 (1-291) x AR274513\_03 (1-110000)

Qy	2	LysLysValIlelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu	21
Db	79512	AAGTCGTCAATTATTCGAGGTAATGGAACAAGTTTAAATCAATGACTATAGTTTATTA	79453
Qy	22	ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu	41
Db	79452	CCTAAGAATTATGATGTTTCCGTTGCAATCAATTTATTTTGAAGATCATATTATTTCTT	79393
Qy	42	GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr	61
Db	79392	GGCAAGAAATAAAAAAGATTTTTTAATGTTCTACAAATTTTGGACATATCTACT	79333
Qy	62	LeuLysHisLeulleGlnAsnGlnGluTyrGlu---ThrGluLeulleMetCysSerAsn	80
Db	79332	TTTATGCAATTAAATTAATAATGAATATGAATATGCTGATATTAATTTATCATCTTTT	79273
Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAspTyrPhePro	100
Db	79272	GTCAATTTAGGAGATTGAGATTTTAAAGAAA---ATTAAAAATGTACAAAATTTACTACA	79216
Qy	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120
Db	79215	CAAGTTGATATGAGCATTTATTTTAAACAAGTACCCGCCCTTTGATGCTCTATTACAA	79156
Qy	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140
Db	79155	TATAACGAATTATATGAATAAAGAGAATTACATCAGCGGTTTATATGTGTGAGTGCA	79096
Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGly	159
Db	79095	ACTGTAATGGGTATAAGATCTTTATTTTAAACAGGTATGATTTTATCAAGAAAAAGGG	79036
Qy	160	SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys	179
Db	79035	AATCCTTACGCAATTCATCAAAAAGAAATATTAATTAATTTATTAATACCTCTCTTTTCA	78976



LOCUS	CQ872844	4277 bp	DNA	linear	PAT 27-SEP-2004
DEFINITION	Sequence 457 from Patent WO2004078949.				
ACCESSION	CQ872844				
VERSION	CQ872844.1				
KEYWORDS	GI:52746651				
SOURCE	Haemophilus influenzae				
ORGANISM	Haemophilus influenzae				
REFERENCE	Bakteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.				
AUTHORS	1 Bakalez, L.O., Munson, R.S. and Dyer, D.W.				
TITLE	Genes of an otitis media isolate of nontypeable haemophilus influenzae				
JOURNAL	Patent: WO 2004078949-A 457 16-SEP-2004; CHILDREN'S HOSPITAL, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..4277				
	/organism="Haemophilus influenzae"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:727"				
misc_feature	2280				
	/note="n = a, c, g, or t"				
misc_feature	2283				
	/note="n = a, c, g, or t"				
misc_feature	2288				
	/note="n = a, c, g, or t"				
ORIGIN	.				
Alignment Scores:	6.97e-78 Length: 4277				
Pred. No.:	692.50 Matches: 143				
Score:	67.5% Conservative: 50				
Percent Similarity:	50.0% Mismatches: 84				
Best Local Similarity:	44.3% Indels: 10				
Query Match:	2 Gaps: 6				
DB:	US-10-734-719-9 (1-291) x CQ872844 (1-4277)				
QY	2	LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu	21		
DB	868	AAGTCGTGTCATTATTGCAGGTAATGGAACAAGTTTAAAAATCAATGACTATAGTTATTA	809		
QY	22	ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu	41		
DB	808	CCTAAAGATTATGATGTTTCCGTTGCAATCAATTTATTTTCGAGGATCATTAATTTCTT	749		
QY	42	GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr	61		
DB	748	GGTAAGAAAAATAAAAAAGGTATTTTAAATGCTGTAATTTTGAACAATACTATACG	689		
QY	62	LeuLysHisLeuIleGlnAsnGlnGluTyrGlu---ThrGluLeuIleMetCysSerAsn	80		
DB	688	TTTATGCAATTAATAAAATAATGAATATGAATATGCTGATATATTTTATCATCTTTT	629		
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100		
DB	628	CTAAATTTAGGGGATTCAGATTAAAGAAA---ATCCAGCGTTTAGAAAAATTAACCA	572		
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120		
DB	571	CAAAATCGATCTTGTCATAGCTATTTAAAAAACTACGAGCTTTTGATGCTCATTTACAA	512		
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140		
DB	511	TATCAGCAACTATATGAGAAATAGAGGATTACATCAGGCGTCTATATGTGTGAGTGCA	452		
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGly	159		
DB	451	ACTGCAATGGGTATATAAGATCTTTATTTGACAGGCAATTCATTTTATCAAGAAAAAGG	392		
QY	160	SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys	179		
DB	391	AATCCTTAGCATTTTCATCATCAACAGAAAAATATATTAATATTATCTCTTTTCA	332		

LOCUS	CQ872844	4277 bp	DNA	linear	PAT 27-SEP-2004
DEFINITION	Sequence 457 from Patent WO2004078949.				
ACCESSION	CQ872844				
VERSION	CQ872844.1				
KEYWORDS	GI:52746651				
SOURCE	Haemophilus influenzae				
ORGANISM	Haemophilus influenzae				
REFERENCE	Bakteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.				
AUTHORS	1 Bakalez, L.O., Munson, R.S. and Dyer, D.W.				
TITLE	Genes of an otitis media isolate of nontypeable haemophilus influenzae				
JOURNAL	Patent: WO 2004078949-A 457 16-SEP-2004; CHILDREN'S HOSPITAL, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..4277				
	/organism="Haemophilus influenzae"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:727"				
misc_feature	2280				
	/note="n = a, c, g, or t"				
misc_feature	2283				
	/note="n = a, c, g, or t"				
misc_feature	2288				
	/note="n = a, c, g, or t"				
ORIGIN	.				
Alignment Scores:	6.97e-78 Length: 4277				
Pred. No.:	692.50 Matches: 143				
Score:	67.5% Conservative: 50				
Percent Similarity:	50.0% Mismatches: 84				
Best Local Similarity:	44.3% Indels: 10				
Query Match:	2 Gaps: 6				
DB:	US-10-734-719-9 (1-291) x CQ872844 (1-4277)				
QY	2	LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu	21		
DB	868	AAGTCGTGTCATTATTGCAGGTAATGGAACAAGTTTAAAAATCAATGACTATAGTTATTA	809		
QY	22	ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu	41		
DB	808	CCTAAAGATTATGATGTTTCCGTTGCAATCAATTTATTTTCGAGGATCATTAATTTCTT	749		
QY	42	GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThr	61		
DB	748	GGTAAGAAAAATAAAAAAGGTATTTTAAATGCTGTAATTTTGAACAATACTATACG	689		
QY	62	LeuLysHisLeuIleGlnAsnGlnGluTyrGlu---ThrGluLeuIleMetCysSerAsn	80		
DB	688	TTTATGCAATTAATAAAATAATGAATATGAATATGCTGATATATTTTATCATCTTTT	629		
QY	81	TyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro	100		
DB	628	CTAAATTTAGGGGATTCAGATTAAAGAAA---ATCCAGCGTTTAGAAAAATTAACCA	572		
QY	101	AspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLys	120		
DB	571	CAAAATCGATCTTGTCATAGCTATTTAAAAAACTACGAGCTTTTGATGCTCATTTACAA	512		
QY	121	PheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAla	140		
DB	511	TATCAGCAACTATATGAGAAATAGAGGATTACATCAGGCGTCTATATGTGTGAGTGCA	452		
QY	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGly	159		
DB	451	ACTGCAATGGGTATATAAGATCTTTATTTGACAGGCAATTCATTTTATCAAGAAAAAGG	392		
QY	160	SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys	179		
DB	391	AATCCTTAGCATTTTCATCATCAACAGAAAAATATATTAATATTATCTCTTTTCA	332		

QY 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199  
 Db 331 CAAATAAAAGTCAAGCGATATCCATCTATGGATATGATTTAAATGCATTTATTT 272  
 QY 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219  
 Db 271 TTACAAAACATTATGGAGTAATATTATTATGCAATTCGCCAGAAAGCTCTATGTAAT 212  
 QY 220 PheIleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsn 237  
 Db 211 TATTTCCTTTATCACCA---CTGAATAACCAATTTACTTTTATTCTCGAAGAAAGAAA 155  
 QY 238 AsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsn 257  
 Db 154 AATTACACCAAGATATTATTATTCGCCGAAGTTTGATATATAAAATAATCGT----- 101  
 QY 258 IleAsnPheLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu 277  
 Db 100 ---ATATATTCCAACCAAGAAATTTACCAAAATCTGTTTTCGG-TTGATCTGGGATATA 45  
 QY 278 LeuArgLeuProSerAsp 283  
 Db 44 TTACGTTTACCTAAGAAAT 27

RESULT 91  
 AE006157/c  
 LOCUS  
 DEFINITION Pasteurella multocida subsp. multocida str. Pm70 linear BCT 16-JUN-2004  
 of the complete genome.  
 ACCESSION AE006157 AE004439  
 VERSION AE006157.1 GI:12721516  
 KEYWORDS  
 SOURCE Pasteurella multocida subsp. multocida str. Pm70  
 ORGANISM Pasteurella multocida subsp. multocida str. Pm70  
 Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 Pasteurellaceae; Pasteurella.  
 1 (bases 1 to 10024)  
 May,B.-J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and  
 Kapur,V.  
 Complete genomic sequence of Pasteurella multocida, Pm70  
 Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)  
 11248100  
 2 (bases 1 to 10024)  
 Zhang,Q. and Kapur,V.  
 Direct Submission  
 Submitted (24-OCT-2000) Department of Veterinary Pathobiology,  
 University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN  
 55108, USA

FEATURES  
 source  
 1..10024  
 /organism="Pasteurella multocida subsp. multocida str.  
 Pm70"  
 /mol\_type="genomic DNA"  
 /strain="PM70"  
 /sub\_species="multocida"  
 /db\_xref="taxon:272843"  
 complement(79..1470)  
 /gene="PM1167"  
 complement(79..1470)  
 /gene="PM1167"  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AAK03251.1"  
 /db\_xref="GI:12721517"  
 /translation="MNLAEPTKSGFNRAIIFLDIVIFILNLVLPDANKGLA  
 LVLVIAVLMTEALHVTVALLVPLAIGLVATKNLAAPADTIFLFFGGALAT  
 ALHTQKLDRTAKMALARGKLFVASIYLFSTAFLSMMNSNTATAMMPLAMGLV  
 SMDREAHNTYVYFILLGAYSANIGMGTLVGSPPNAIVASQLKLTFSDWLQSGPIN  
 QMLMPDMLLVFVKPKLVNREKHEFVLEMKSRITLICIFVTILCWVFSQIN  
 PMLSNLGLKXKMSFDSVALLAAVVICSTGVATWKQIQNTDVGWMLFGGGLTIS  
 AVLKDSGASKLADGIVFLIEGGHFYLLGLLVAAPFIIFLTFTSNTASAALLVPFIS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

IAOSLGMPGLGIALIIGLGASCAFMPLPVATPPNAIVFGTGEIRQSEMVVRVGVLLNIVC  
 IFVIATVGVLPWL"  
 complement(1688..2188)  
 /gene="menG"  
 /note="synonym: PM1168"  
 complement(1688..2188)  
 /gene="menG"  
 /codon\_start=1  
 /transl\_table=11  
 /product="MenG"  
 /protein\_id="AAK03252.1"  
 /db\_xref="GI:12721518"  
 /translation="MYIDTSELCDIYLDQDVVVEPIFSFSGGVNRFYGVKVTTVKCFEN  
 NGLIIDILEENGEGRVLLIDGGAVRRALVDAELAQLAVDNGWEGIIYGAVRQIQQL  
 BEMDIGIHALAPIVCGADKQNGEIDVPVNFVGGVTFPPEDVYVADLTGIIILSQPELEL  
 DEFENDA"  
 complement(2239..3150)  
 /gene="menA"  
 /note="synonym: PM1169"  
 complement(2239..3150)  
 /gene="menA"  
 /codon\_start=1  
 /transl\_table=11  
 /product="MenA"  
 /protein\_id="AAK03253.1"  
 /db\_xref="GI:12721519"  
 /translation="MEWTNNVLKKVETARPRTLPLAIASIITCASLAFWSKFESWTI  
 TWCLITTVLLQILSNFANDYGHQKSGDTADRIGPLRAIQOGEISASQLKKGILVVA  
 LASLCSGAILIIVAYOSLADLAFSLGLLAAVAAITTYGAKPYGYMGGLSLVLF  
 FGLGVGGSYLTQQLNLAILLPAFASGLLASAVLNINLRDIEQDSQVGNKTLAVR  
 LGSEKRIYHACLLIVAILCYLIFTLCYVQKLTGTFILFALPLFVKHALFVYQNKSP  
 QLRLPMAQMSLLALLINLLFSLGLLIS"  
 3246..3986  
 /gene="PM1170"  
 3246..3986  
 /gene="PM1170"  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AAK03254.1"  
 /db\_xref="GI:12721520"  
 /translation="NSDLSLQLHAIGIHTPYKEKFSVPQPNLVQDGTGILELLPPY  
 NOAETVRGLAOFSLHWLIFQEDRVATGKWRPTVRPRLGNORVGVFASRSTHRPNPL  
 GLSKVELRVEQCNQKVLHLGNADVLDVDTPIFDIKPYLAYADSEPEAKSGFAOEKPE  
 CTLQVIFSEQAQNALQKIEKRPHFRFITEVIAQDPRPAYQKMSLSERVYGIRLHEF  
 NTRWKEATEEQARILDIIEVEKKCD"  
 complement(4011..4385)  
 /gene="PM1171"  
 complement(4011..4385)  
 /gene="PM1171"  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AAK03255.1"  
 /db\_xref="GI:12721521"  
 /translation="MMKLLIALASIIAISGMAVAKGFGQNNNNAGCGFNPAGINSV  
 RAALNAQDDMPVSIIEGYIVRQIDGDEFIFRDVAGSEIRIDVSRAMWGQTIQPNDLII  
 IQGKVDKWNKTDIDVKNIIKK"  
 4580..5425  
 /gene="PM1172"  
 4580..5425  
 /gene="PM1172"  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AAK03256.1"  
 /db\_xref="GI:12721522"  
 /translation="MLNKIDSMIIQHPQOHLTGIRKFSIEFLFGLKEIRACLFAGL  
 PFIAFITPKAGFLGPLRYDVLIIIFAILLQVINFYTKFTLDEIKSITLPHLIQFALE  
 LFKTSISSQWSYDPFAYSKIYGVPLTFGMVAAGSYIIQAWRLFDLKI KSHPPFL  
 STLTAILIYNFETHYIGDYRWYLAVFALGVARTIVYFTPYDKPRKMPLLFAFMLI  
 GFPIWLAENWGTFIGIWRYPNQIGAMSVHVSWSAWALLIVMTFTIVANLKHKKSI  
 RVSRD"







FFORGVIOAIGFLFLFLYKFGDSFATTLQTKFIYDMGFSKEDIAIIVKSKLSWSSI  
LSLGGAGMLKLGINRALWGLVQMTVIGGFIWLSAFGHFDVITSALWKLGVVIA  
ASYIGVGLGTAFAVMARESNPLYTATQALFISLSALPSKVLGILSGYVVGAVGY  
QYWFCLFLAIPGMLCLFWFVAPWKQKNKTSV"

gene

/gene="adk"

/db\_xref="GI:43591"

/db\_xref="GOA:P24323"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

/db\_xref="UniProtKB/Swiss-Prot:P24323"

/translation="MKIIILGAPAGKGTQAOQFIMNKGFIQIISTGDMFRAATKAGTE  
LGQKALMDEGLVDELTAVALKDRIQAQDANGFLDLPFRPTIPODALDKSGVK  
IDVLEFDVDEVRMGRVHOAGSRSHIVYVPPKVGKDDVTGEDLIIRADDK  
PETVLDRLAVYHKQTSPLDIYQAEAKAGNTQYFRDGTQKVEVSGELDKLG"

ORIGIN

Alignment Scores:

Pred. No.: 1.14e-68 Length: 4977  
Score: 621.00 Matches: 129  
Percent Similarity: 59.7% Conservative: 46  
Best Local Similarity: 44.0% Mismatches: 64  
Query Match: 39.7% Indels: 54  
DB: 15 Gaps: 6

US-10-734-719-9 (1-291) x HILIC3 (1-4977)

QY 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21  
DB 588 AAGTGTGTCATTATTCAGGTAATGGAACAAGTTTAAATCAATGACTAGTTATTA 647  
QY 22 ProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrLeu 41  
DB 648 CCTAAGATTATGATGTTTCGGTTCGCAATCAATTTATTTTGAAGATCATATTTCCT 707  
QY 42 GlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrThr 61  
DB 708 GGCAAGAAATTTAAATGTA----- 728  
QY 62 LeuLysHisLeuGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyr 81  
DB 729 -----CAAAATTAACAACAAGTT----- 749  
QY 82 AsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAsp 101  
DB 749 ----- 749  
QY 102 AlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPhe 121  
DB 750 ---GATATTGGACATTATTATTAACAAGCTACCCGCTTTCGATGCTATTTACAAATAT 806  
QY 122 HisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIle 141  
DB 807 AACCAATTATGATAAATAAGAGAATTACATCAGCGCTTTATATGTCGAGTGGCAACT 866  
QY 142 AlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGln---AsnGlySer 160  
DB 867 GTATGGGTTATAAGATCTTTATTTAACAGGATTATTTTATCAAGAAAAGGGAAT 926  
QY 161 SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn 180  
DB 927 CCTAGCATTTTCATCATCAAAAGAAATATTATTAATTTACCTCTTTTTCACAA 986  
QY 181 AspArgSerHisTyrIleGlyHisSerLysAsnThrAspLysAlaLeuGluPheLeu 200  
DB 987 AATAAAGTCAAAAGCGATATCCATCTTATGGAATATGATTAAATGACATTTATTTT 1046  
QY 201 GluLysThrTyrLysLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe 220  
DB 220 ----- 220

Db 1047 CAAAACATTATGAGTAAATATTTATGCAATTCGCCAGAAAGTCCTCTATGTAATAT 1106  
QY 221 IleGluLeuAlaProAsnLeuAsnSer-----AsnPheIleIleGlnGluLysAsnAsn 238  
Db 1107 TTTCTCTTATCACCA---CTGAATAACCAATTTACTTTTATTTCTCGAAGAAAAGAAAT 1163  
QY 239 TyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIle 258  
Db 1164 TACACACAGATATTTAAATTCGCCGAAGTTTGATATATAAAAAAATTCGT----- 1214  
QY 259 AsnPhelLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 278  
Db 1215 ATATATTCCAAACCAAGAAATTTACAAAATCTGATTTTTCGTTGATCTGGGATATTA 1274  
QY 279 ArgLeuProSerAspIleLysHisTyrPheLysGlyLys 291  
Db 1275 CGTTTACTTAATGATATAAAAACACGCCTTAAATCAAGA 1313

RESULT 94

AF486554

LOCUS

DEFINITION

AF486554 9515 bp DNA linear BCT 08-FEB-2005  
Campylobacter jejuni strain RM2227 HsdR (hsdR) gene, hsdR-1 allele,  
complete cds; RloD (rloD) gene, complete cds; HsdS (hsdS) gene,  
hsdS-6 allele, complete cds; MloA (mloA) gene, complete cds; and  
HsdM (hsdM) gene, hsdM-1 allele, complete cds.

ACCESSION

AF486554

VERSION

AF486554.1

KEYWORDS

GI:19881258

SOURCE

Campylobacter jejuni

ORGANISM

Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Campylobacteraceae; Campylobacter.  
1 (bases 1 to 9515)  
Miller, W.G., Pearson, B.M., Wells, J.M., Parker, C.T., Kapitnov, V.V.  
and Mandrell, R.E.

TITLE

Diversity within the Campylobacter jejuni type I

restriction-modification loci

JOURNAL

Microbiology (Reading, Engl.) 151 (PT 2), 337-351 (2005)

PUBMED

15699185

REFERENCE

2 (bases 1 to 9515)

Miller, W.G., Pearson, B.M., Wells, J.M. and Mandrell, R.E.

Direct Submission

TITLE

Submitted (21-FEB-2002) Produce Safety and Microbiology, USDA, ARS,

JOURNAL

WRRRC, 800 Buchanan St., Albany, CA 94710, USA

FEATURES

Location/Qualifiers

1..9515

/organism="Campylobacter jejuni"

/mol\_type="genomic DNA"

/strain="RM2227"

/db\_xref="taxon:197"

75..2384

/gene="hsdR"

/allele="1"

75..2384

/gene="hsdR"

/allele="1"

/note="putative type I restriction subunit"

/codon\_start=1

/transl\_table=11

/product="HsdR"

/protein\_id="AAM00865.1"

/db\_xref="GI:19881259"

/translation="MQNLYEDDTRVKFIDAKLYDSSWNEENIKRNYFTDGRKLGG  
KRAQRKADYLLRYEGNNLAIIEAKFKSDPLDGLSGIEYAKILNVPVSSNGBK  
YEYDIRHSGKYIDKFPSPKELFDRIYGNVKEWYKLLTQNYMYIPQKELRYQKIAI  
DKVTEAILNNKNNRILLTATGTTTIAFNLCYRLRLEARNKWKDQKILPLCDRV  
SLRQALGFENPIESDCKAIISAEIKKNDGKIIITNANIFFGIYQSLAANSKEQEQE  
EQESKFQYKDFDLIIIECHRGANEESWRVLEYFYATQGLTATPKKEEN  
IDTYEFGESYDLSGIEEDFLTPYKVKLIKTLTSDGYTNPDDLQGELEKGY  
KQSFERNIFLPNYNDFAKLILELINPMDKTIIFCVAQAHANEVRAIDKYKVRD  
DVCYRVDEGKILGLDYKLFQDNKSPVILTSSKMLTTGVDANKVRNIVLLANGS  
MVFEKQIIGRTRVVEGKDFTLDFVGTATKLFYDPKWDGEKIEELKEQDEKEKIK  
HIKQTKESKESKSVTHLKGTKLVLDITTYVGAQKPLSTKTFLEFLIGLGEYV

gene	DDEAKLRISWQKNRERFLKALANDGVDEDAKDRLREIFQKDCDIYDLAHLSPNAE	
	IKTRQERVLOVENGSEFLKRFQKEKAIKFIEFLGRYQVQGIKDFDGLKPLIELSSLG	
CDS	NVRELADEFGGLEILKQSFDDLQREIYLN"	
	2577..5621	
	/gene="r10d"	
	2577..5621	
	/gene="r10d"	
	/codon_start=1	
	/transl_table=1	
	/product="r10d"	
	/protein_id="AA000866.1"	
	/db_xref="GI:19881260"	
	/translation="MFGKLFSLNFGSIKQENHNHNSNIISFIKQNEDYFYNGEFL	
	KSPFELKEVKNDSKKNVLLVNEAKYFDLCNKKTKENLYLLEKEYKNFIDIS	
	PKETQLSLCMHEKDPKNEIKQYFLIEKQNRSENYDFPMVALNTGCIKQAKKLPDK	
	LKEKESEFLKANLYAQFFKEQENEDALLFIELCETLIQDNKLNFKQKLIILETYE	
	IEKFTFRYNIISLNKNYQIDILENIIKNDNLQYFGDYQRYKINYCEILLFS	
	BENDKFIYMNKDELNFHVFYSIDIVKNDIDHNIQDKILETKDFKPMYDIITYC	
	DIHKDKAHSFLSSNYKLLDEHNLIFYIKICLDKKFELTOEIKHFIAKSESNIITYC	
	FLNLKLIASNDELEKLYSKLIDITLSECFIKOILFQEQNFNSWTDILKFKDKYN	
	GLVELALSIFYNNININQFENFKINDQTLKCEFISSIYLCNGEYKSVKIQDUWSD	
	KNLDKTLKSHTLAILKQYDDNKEILDNLTKAINLYKADLDNLKLIETQLLYCS	
	LLFYKIDTDIIFLKNQLLECPINNLQDETVCNICEIYNKYITNNNIVQDTLIYDKN	
	TGOAVISKYNKDPISYNIISIDSPFDLKDNDYIKYSIFPDIVCTIISQKNLP	
	NVSKINDNNPLPEFRSEVNNITPPIEQERKLTWQFNQNCYLFHLKNSLGLDYIE	
	LIPMLIEDINFNHNCIPYGGSKILTFSSIVLQHLKNDLVLRQNIPIPKSIIL	
	EIQNKQIENSKNKEIYLSNGLSKIDNSKVKKEKLEHLKLVDLKQIIDDTKYPIIS	
	NLEDPISLKEIEIGFISCCINQYQIITEDRFPGLABEINAPQIVSNAMSLIIEKD	
	YIDRAILDSKKYFVFGYWAYINATCLYKNISYNISSHDKKIIQIINKYGLDK	
	LKQYKNTVEVLYPKNFPKNDYVVKHNIPELLNLLEENDNAK"	
gene	5605..6798	
	/gene="hsds"	
CDS	/allele="6"	
	5605..6798	
	/gene="hsds"	
	/allele="6"	
	/note="putative type I specificity subunit"	
	/codon_start=1	
	/transl_table=1	
	/product="Hsds"	
	/protein_id="AA000867.1"	
	/db_xref="GI:19881261"	
	/translation="MIMQNKTNLPQGWKVTLSIEIIVTGSTPSKSNLDPYKDPY	
	FKFQSPDFQGVLENAGNLSKLPDKARQPPKTLIVVCISGLKVALTVIGSCNQ	
	QINAIIPHNITIAEYIYYICSSKFSQILFSKAPQTTAILNKTEFSKLEIYPKDIK	
	EQEIRVILDESFNADIESKILEODLNLDELMOGLKAFNPLKDNKAKENYKLPQG	
	WEWKSLEISNISAGGDKPKNCTESKTKNQIPVYANGVNNGLVGYTDKATIIKPS	
	LTIISARGTIGVCIRKEPYFPVIRLLISLPCENILCLHLYFLCLNFFAKGEGSSIPQ	
	LTIIPKFSLIQIPLPKLEQEQIAGHLDPFVEFKALKELYTKELKDYEEELKQSLUNKA	
	FKGEL"	
gene	6918..7988	
	/gene="mloA"	
CDS	6918..7988	
	/gene="mloA"	
	/codon_start=1	
	/transl_table=1	
	/product="MloA"	
	/protein_id="AA000868.1"	
	/db_xref="GI:19881262"	
	/translation="MKEIIPQKLPINIELNANIYSLIIKASRKLAEINGLSKSINPNS	
	ILINALIQEARLQKSEIENIITTHDELFLSRVDENKLTFRAAKVKDYSALKKGYELL	
	KKEQLLRNHLIEQRLERNNAGFRQSGTMLKNLPITGEIKHIPQPNPNNIOELMGN	
	LELYINDSLDFPLVIAIHYQFESHPFDGNDRTGRIINILYLVYKGLDLPFI	
	LYLSAYIVKSDVEYKLVQVRDEGDIKWEIYILKIEOTAVTETIETIIKIEKMNSN	
	VGEILQNKTPYSKDFVELLSHPYTKIDFLIKLIDISRQSAKYLKICENLGILECI	
	KMGKSNYYINIELLKLFRKGF"	
gene	7989..9473	
	/gene="hsdm"	
CDS	/allele="1"	
	7989..9473	
	/gene="hsdm"	
	/allele="1"	
	/note="putative type I modification subunit"	

/codon_start=1	
/transl_table=11	
/product="HsdM"	
/protein_id="AA000869.1"	
/db_xref="GI:19881263"	
/translation="MONKIDKITDILRRDDGISGAMHYTEQIGWILFLKELDDYETVL	
KDLAFLDGDKYKSILOQEFKFSWVAAPKDKLDVKNALSGDLELFEVNFELPYLKN	
FKNDDFKSIYKIGIIEFIDNRNIAHGHTLRVINIIDIESFNKEDVEFALGEVYEK	
LLKDMGSDGNSGEFVTPRLIKTMVEVIDPKPERYIDPSCGCGFLVSFLHLILYK	
LTKGKKNANLSVEELEFLKNDALFGKEKTPLSKVGMMMLHEIISPNLIKNTWLTK	
KITDITEKDYEVILANPPFGKKEQIQENFFIKSNATELLFLQHLKLSLKNNGRCA	
IIVPGVLVFNQSNAPFVSKDLDLDFNLVSLPGSVGLFPYSAVNTVNLVFFSGKKC	
ICEGQVYVYELIPYKLTNKPLEYAHFKFELCKYKERRIKTANSNLVLSVKELBERNY	
DLSAKNPVKEKILRTSEILNSLENLKTQKYLNELKSLIK"	
ORIGIN	
Alignment Scores:	
Pred. No.:	Length: 9515
Score:	Matches: 82
Percent Similarity:	Conservative: 46
Best Local Similarity:	Mismatches: 112
Query Match:	Indels: 113
DB:	Gaps: 18
US-10-734-719-9 (1-291) x AF486554 (1-9515)	
QY	12 SerLeuLysGluLeuAspTyrSer-----ArgLeuProAsnAspPheAsp 26
Db	2880 TCTTTTAAAGAACTCAATTATCATTTGTGTATGTCATGAAAAGATCCAAATAATTAAC 2939
QY	27 ValPhe-----ArgCysAsnGlnPheTyrPhe 35
Db	2940 GAGATAAAACAATATTTTAAATAGAAAAACAATAAGATCTAATGAG---TAITTT 2996
QY	36 GluAspLysTyrTyrLeuGly-----LysLysCysLysAlaValPhe----- 49
Db	2997 GATTTTATGTCATTAATAACAGGAGATATAAAACAACAAAAAATTTATTGACAAG 3056
QY	50 -----TyrThrProAsnPhePhe 55
Db	3057 TTAAAAAGAAAAAGAGAGTGAATTTTAAAAAGCAAAATTTATACGCACCAATCTTTTTT 3116
QY	56 PheGluGln----- 58
Db	3117 AAAGAACAAAAATGCAATATGTTTATGAAATATGTGAAACACATTATACAA 3176
QY	59 -----TyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGlu 74
Db	3177 GACAATAAATTAATTTTTCGAGAAAAAAATTTATTAGAACGCTGTATGAAATTGAA 3236
QY	75 LeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThr 94
Db	3237 AAATTTTCTAGAAAAATACAAATAAGTATTTTAAAAATAAAAACTATATAAAAAAT 3296
QY	95 PheTyrAspTyrPheProAsp-----AlaHisLeuGlyTyrAsp 107
Db	3297 TATCAAGATATCTTGGAAAAATATATAAAAAATGCACAATCTGCAATATTTGGGTTTGT 3356
QY	108 PhePheLysGlnLeuLys-----GluPheAsnAlaTyrPhe 119
Db	3357 TATCAAAAGTATATAAATATAATTAATTTGTGAAATTTTATTATCAATTTGAAGAAAAATGAT 3416
QY	120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal 139
Db	3417 AAATTTATAGATTTTACATGAATAATAAAGTGAACCT----- 3455
QY	140 AlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159
Db	3456 -----TTTAAACCATTTTATTTTTCATATATATGATTTGTAATAAAAAATGAT 3500
QY	160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLys 179
Db	3501 ATAGATCAC-----AATATAATACAAAGATAAAATTTTAGAA---ACTAAGATTTTAAA 3551

```

Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 3552 TTAATGTATGACTATATACATATATGTGATATCATTAAGATAAATAAGGACATAGCTTC 3611
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 3612 TTATCTAGCAATTATAAATTA-----TTAGATGAGCATAACTTAATATTTTAT 3659
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysPheSerLys 239
Db 3660 TTATATAAAATTTGCT-----CTCGATAAGAAATTTGAAATAACGACGAA----- 3704
Qy 240 ThrLysAspIleLeuIleProSerSerGluAla-----TyrGlyLysPheSerLys 256
Db 3705 ATAAACATTTTATAGCAATAAATCAGATCCAAACATAGTGAT-----TATTTTATA 3758
Qy 257 AsnIleAsnPheLysLysIleLysIleLys-----GluAsnValTyrTyrLysLeu 273
Db 3759 AATTAAATAAATAAATAAATAGCTTCAATCATGAATAGAAAATTTATATCTTAAGCTA 3818
Qy 274 -----IleLysAspLeuLeu 278
Db 3819 GACATAACTTTAAGTTCGAGTTTATAAAGATATAATA 3857

```

```

RESULT 95
LOCUS AF486556
DEFINITION Campylobacter jejuni strain RM2240 HsdR (hsdR) gene, hsdR-1 allele, complete cds; rloD (rloD) gene, complete cds; HsdS (hsdS) gene, hsdS-3 allele, complete cds; MloA (mloA) gene, complete cds; and HsdM (hsdM) gene, hsdM-1 allele, complete cds.
ACCESSION AF486556
VERSION AF486556.1 GI:19881270
KEYWORDS Campylobacter jejuni
SOURCE Campylobacter jejuni
ORGANISM Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.
REFERENCE 1 (bases 1 to 9561)
AUTHORS Miller, W.G., Pearson, B.M., Wells, J.M., Parker, C.T., KapitonoV, V.V. and Mandrell, R.E.
TITLE Diversity within the Campylobacter jejuni type I restriction-modification loci
JOURNAL Microbiology (Reading, Engl.) 151 (PT 2), 337-351 (2005)
PUBMED 15699185
REFERENCE 2 (bases 1 to 9561)
AUTHORS Miller, W.G., Pearson, B.M., Wells, J.M. and Mandrell, R.E.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Produce Safety and Microbiology, USDA, ARS, WRRC, 800 Buchanan St., Albany, CA 94710, USA
FEATURES
source
1. .9561
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="RM2240"
/db_xref="taxon:197"
61. .2370
/gene="hsdR"
/allele="1"
61. .2370
/gene="hsdR"
/allele="1"
/notes="putative type I restriction subunit"
/codon_start=1
/product="HsdR"
/transl_table=11
/protein_id="AA000875.1"
/db_xref="GI:19881271"

```

```

Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 3552 TTAATGTATGACTATATACATATATGTGATATCATTAAGATAAATAAGGACATAGCTTC 3611
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 3612 TTATCTAGCAATTATAAATTA-----TTAGATGAGCATAACTTAATATTTTAT 3659
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysPheSerLys 239
Db 3660 TTATATAAAATTTGCT-----CTCGATAAGAAATTTGAAATAACGACGAA----- 3704
Qy 240 ThrLysAspIleLeuIleProSerSerGluAla-----TyrGlyLysPheSerLys 256
Db 3705 ATAAACATTTTATAGCAATAAATCAGATCCAAACATAGTGAT-----TATTTTATA 3758
Qy 257 AsnIleAsnPheLysLysIleLysIleLys-----GluAsnValTyrTyrLysLeu 273
Db 3759 AATTAAATAAATAAATAAATAGCTTCAATCATGAATAGAAAATTTATATCTTAAGCTA 3818
Qy 274 -----IleLysAspLeuLeu 278
Db 3819 GACATAACTTTAAGTTCGAGTTTATAAAGATATAATA 3857

```

```

SLRDAQLEFNPIESDCKAISAEIKNKDGKIITNANIFFGIYQSLAANSKEQENIQE
EQSKFYLOVPKDFDLIIIDCHRGANEEGSRWAVEYFSYATQDGLIATPKEEN
IDTYVEGENVDYSLKSGEDGFLTPVKVKLIKTLLSDGYTNPDDLIQGLEKGFY
KQSEFERNIFLPNYNDPIAKKILSELNPMKDTIIFCVNOAHANEVYKRAIDYKVRD
DYCVRVSDGKIGLDYIKLGFQDNDKSYPIVLTYSKMLTTGVDAKNRNVILLANIGS
MVPEKQIIGRTRVYEGKDFITILDFVAGATKFDYDPKWDGEKIEELKQDEKEKIKKE
HIKQREESKEKSVTHLTKLVLDITTYTGAQGPLSKTEFLFLKGLDPTIYV
DDEAKREIWSDDKRRERFLKALANDGVDEDALDRLREIFQKDCDIYDLAHLSPNAE
IKTRQERVLOVENGELFKRFQEKAIKIFBFLGRYQQVGIKDFDGLKPLIELSLG
NVRLEDBFGGLEILKQSPFDLQREIYLN"
2563. .5607
/gene="rloD"
2563. .5607
/gene="rloD"
/codon_start=1
/transl_table=11
/product="rloD"
/protein_id="AA000876.1"
/db_xref="GI:19881272"
/translation="MFGFKLPSFLNFGSIKOENHNHNSNIISFIKQEDYFYNGEKF
KSEFILKEYEDNSDKKNYLLAVNEAKYFDLCNYKTKTENLYLKEYKNFIDIS
FKETQLSLCMHEKDPKNEIKOYFLIEKQTRNSNEYDFPMYALNTGDIKQAKKLPDK
LKEKSEFLKANIYAOSFFKEQENEDALLFIELCETLIQDNKLMFLQKKIILEIYDYE
IEKFTFKYNIISILKNKYIKNYDILENIKNNDLYFGFDYQYIKYNYCEIILSF
EENDKTFEFYMNKDELNFHFYSYDIDVKNDDHNIQDKILETFDKMLDYIITYC
DIHKDKAHSFLSNKYLLDBHNLIFYIKICLQKKFETOBIEKHPFANKSBSNIYVY
FIALKLKIASNDELEKYSKLDITLSCFEDIILFQEQNFNSWTDILAFKQKYN
GLVLEALSIFYNNINOFNPIKNIDOTLYQKEISSIYLNCGEYKSVKITQDLMSD
KNLDTLSLHRLAILQYQYDDNKEILDNDLTKAINYLKADLNKLEILLQLYCS
LLFYKIDTDIIFKLNKQLLECPNNLDQETVNCICEYKNYITNNIVQDTILYDKN
TQYVIGSKYNNKIDPSYISIIIDSEFDLKLKDDNYIKYSIFDIYCTIISQKNLP
NGVININDNNNPLEFFREVNITFPIBOERKLIWOFNNONCYLFLKNSLGLDYIE
LIPLMIEDEDINFHNNCIPYSGSKILTFSIVLLOHINKLDIVLQKNIFPKYSIIL
EIQNKIONIENSKNKEIYSSNGLSKIDNSKVEKLEYLEIKLVDLKQIIDDTKPIIS
NFKNISIDLKQEIINFSTCCINQOLITEDRFFGYLAEBEINAPQIVSNAMSLLIBEK
YIDKAILLDSKKFYVFGYVAYINATCLYKNISSYNIISNDDKKIIQINKYGFLLDK
LKQYKNTYEVLYPKINFPKNDYVYKHNIELLNLTENDNAK"
5591. .6844
/gene="hsdS"
5591. .6844
/gene="hsdS"
/allele="3"
/allele="3"
/notes="putative type I specificity subunit"
/codon_start=1
/transl_table=11
/product="HsdS"
/protein_id="AA000877.1"
/db_xref="GI:19881273"
/translation="WIMQKNTNLPQGWVKLGEIGELVTGSTPSTKSNLDFYKQDYP
FFKPSDFEQGYFLENAGNLSKLGFDKARQIAPPKTTILVVCIGSLKVALTRVIGSCNQ
QINAIIPHKNIISEYIYICSSKPSILSKAPOTTLAIFNKTSFKLEIYIPKDIK
EQERIVGILDSFAKIDESIILFQDLNLDLMQSLQKAFNPKNAKENYKLPQG
WEKSLGELSNILONGFAASKNNEIPSGYVHLRTHNISTDGNLNDFTIKIKREFIKE
KQSFIEKNDILFNNTNSELVGTALVTQNYNAPSNHLTKIKLKNVNSKLWVYFV
LLKKNYFEKICHQWIGSGINIDKIKIQLPLPLKQEQIQAHLDFVFEKTKALKE
LYTKELKDYELKQSLNNAKFGEL"
6964. .8034
/gene="mloA"
6964. .8034
/gene="mloA"
/codon_start=1
/transl_table=11
/product="MloA"
/protein_id="AA000878.1"
/db_xref="GI:19881274"
/translation="WKKEFIPQKLPNLNIELNANIYSLITKASRKLAEINGLSKSIPNPS
ILINALLOEAKDSSEIENIITTHDELFLSRVDENKLTAAKEVDKESALKKGVELL
KKEQLLRNHAHILEOKRLERNNAGRKQSGTMLKNPITGEIKHIFQPNPNDIQELMGN
LELYINDSDLDFDLVKMAIHYQFESIHFPYDSNGRTGRITNINILYVYKGLDPPFI
LYLSAYIVKNKDEYIKLQKVRDEGDLKWEIYLKQEQATVKTETIETIKIEKMNSN
VGEILOKNTNFYKDFVLLSFHPYTKTDFLIKKLIDISRSQASKYLKICENGLILECI
KMGSRYYINIELLRLFRKGIF"

```





```
Db 3362 TATCAAAAGTATATAATAATATTGTGAAATTTTATTATCATCTTGAAGAAAAATGAT 3421
Qy 120 LysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaVal 139
Db 3422 AAAATTTATAGAAATTTTATCATGATGAATAATAAGATGAACCTT----- 3460
Qy 140 AlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159
Db 3461 -----TTTAAACCATTTTATTTTTCATATATTGATATTGTAAAAAATGAT 3505
Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLys 179
Db 3506 ATAGATCAC-----AATATATACAGATAAAATTTTAGAA--ACTAAAGATTTTAAA 3556
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 3557 TTAATGTATGACTATATTACATATTGTGATATCCATAAAGATAAAAGGCACATAGCTTC 3616
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 3617 TTATCTAGCAATTATAAATTA-----TTAGATGACGATAAATTTAATTTTAT 3664
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyr 239
Db 3665 TTTATAAAATTTGT-----CTCATGAAGAAATTTGAATAAACGAGAA----- 3709
Qy 240 ThrLysAspIleLeuIleProSerSerGluLys-----TyrGlyLysPheSerLys 256
Db 3710 ATAAACACATTTTATAGCAAAATAAATCAGAAATCCACATAGTGTAT-----TATTTT 3763
Qy 257 AsnIleAsnPheLysLysIleLysIleLys-----GluAsnValTyrTyrLysLeu 273
Db 3764 RAATTTAAATTAATAAAATAAGCTTCAAAATGATGAATTAGAAAAATTTATTTCTAAGCTA 3823
Qy 274 -----IleLysAspLeuLeu 278
Db 3824 GACATACTTTAAGTTGGAGTTTATAAAGATATAATA 3862

AC005308 178273 bp DNA linear HTG 12-AUG-2000
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***
AC005308
VERSION AC005308.8 GI:9797719
KEYWORDS HTG; HTGS; PHASE2.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
1 (bases 1 to 178273)
Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
Kurdi,O.B., Conway,A.B. and Davis,R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 178273)
Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
Direct Submission
Submitted (21-JUL-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
On Aug 12, 2000 this sequence version replaced gi:8810450.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as runs of N. However the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 178273: contig of 178273 bp in length.
Location/Qualifiers
```

```
source
1. 178273
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clone="PFYAC492"
/clone="3D7"

ORIGIN
Alignment Scores: 5.45e-05 Length: 178273
Pred. No.: 84
Score: 135.00 Matches: 84
Percent Similarity: 34.7% Conservative: 58
Best Local Similarity: 20.5% Mismatches: 113
Query Match: 8.6% Indels: 154
DB: 12 Gaps: 23

US-10-734-719-9 (1-291) x AC005308 (1-178273)
Qy 2 LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSer---Arg 20
Db 108887 AAAAAAGAAGTGTACGATAATTATGTTGATATAATATATATCTTATATATAGTACAAGG 108946
Qy 21 LeuProAsnAspPhe-----AspValPheArgCysAsnGlnPheTyrPhe 35
Db 108947 AACCCATATAGATAATATACGAGATAAAAGGACAAATTCGTAACACATAATTTATATATT 109006
Qy 36 GluAspLysTyrTyr-----LeuGlyLysLys-----Cys 45
Db 109007 AAAAAAAATATTATTAACACAAATAATAATTAATAAAGAAGAATAATACATCATGT 109066
Qy 46 LysAlaValPhe-----TyrThrProAsnPhePheGluGlnTyrTyrThrLeuLys 63
Db 109067 AGATCTTATTTCGGAATATAATATACGAATAATAATATAAATAAATAAATAAATAAATAA 109126
Qy 64 HisLeuIleGlnAsn-----GlnGluTyr----- 71
Db 109127 TATAACATACATAATGACACAGACACAAATAATTTCTATATACCATAGAAGTATAGATGT 109186
Qy 69 -----GlnGluTyr----- 71
Db 109187 AATGGGAATAACTTTTATATAAAGATTTGAACATTTTCAAGGAAAAATATGTACACAAAT 109246
Qy 72 -----GluThrGluLeuIleMetCys----- 78
Db 109247 GGTCCGATAATAATAATTAATATGTTGTAATTTTAAATAATATATGAATGAATAATGCT 109306
Qy 79 ---SerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLys----- 93
Db 109307 GAGAACACTTAAACCACTTGGATATACAAACATTTAATGATATAAAGGGAAGAATAT 109366
Qy 94 -----ThrPheTyrAspTyrPheProAsp 101
Db 109367 AATATATATCAACCAATGTATGGAATAATAATAATACATATAATAATATATATATATATAT 109426
Qy 102 Ala-----HisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyr--- 118
Db 109427 AAATGTAATAATAATGAATAATATATATACATAAATAAATAAATAAATAAATAAATAAATA 109486
Qy 119 -----PheLysPheHisGlu 123
Db 109487 TTAATAACAAATATATATAATATACATTAACCAATGATGATACAAATTTGGGTATCATAT 109546
Qy 124 IleTyr-----PheAsnGlnArgIleThrSer---GlyValTyr 135
Db 109547 ATTTACGAATGTTTTGTAAACAATTTCTACGACAAAAGATTAACAAGTGTATCTGTTAT 109606
Qy 136 MetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleThrLeuSerGlyIleAspPhe 155
Db 109607 ATG-----CAAAGAGTATATAATAAGTCT----- 109630
Qy 156 TyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAla 175
```



D	b	109631	TATATGAACAAACAGCAATAATATTAAACAAGAATAATAAATCTTTTA-----	109681
Q	y	176	ProAspPheLeuAsnAspArgSerHisTyrIleGlyHisSerLeuAsnThrAspilelys	195
D	b	109682	-----TATTCAATCAGGAATGTTTAATAATCACAGATAAAAAAGAAAATAAAA	109735
Q	y	196	AlaLeuGluPheLeuGluLysThrTyrIlysIleLysLeuTyrCysLeuCysProAsnSer	215
D	b	109736	AGAAATTATGTATATACAGAAATCATATGAAG---AAATTATAT-	109774
Q	y	216	LeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe-----	231
D	b	109775	-----GTTATAAAGTGTAATAATAATTTAAATATCAATAT	109813
Q	y	232	IleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr	251
D	b	109814	ATATATGAAGATTAGATAGATTATTCAC-----ATATCTCTTCAAGT-----TAT	109861
Q	y	252	GlyLysPheSer-----LysAsnIleAsnPheLysLysLysLysLysLys	266
D	b	109862	AAAAAGAAGAACCAATTACAAAATGCGAAAAATGAAATATATAAAAAACTTCTTATT	109921
Q	y	267	GluAsnValTyrTyrIlysLeuIleLysAspLeuLeuArgLeuProSerAspileylshis	286
D	b	109922	AAAAAATTTTCAACACGATATGAACAAAAATATACCGCGCTATTATTTTTTATAAAGCAT	109981
Q	y	287	-----TyrPheLysGlyLys	291
D	b	109982	TTACTACATGATTATTTGTAATAAAAAAG	110008
R	E	S	RESULT 98	
A	E	O	AE014845/c	
L	O	C	LOCUS	
D	E	F	DEFINITION	Plasmodium falciparum 3D7 chromosome 12, section 2 of 9 of the complete sequence.
A	C	S	ACCSSION	AE014845 AE014188
K	E	Y	KEYWORDS	AE014845.1 GI:23496588
S	O	R	SOURCE	Plasmodium falciparum 3D7
R	E	F	REFERENCE	Plasmodium falciparum 3D7
A	A	H	AUTHORS	1 (bases 1 to 250531)
				Barcelon, J., Hall, N., Fung, E., White, O., Bertram, M., Hyman, R.W., Gardner, M.J., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., Carlton, J.M., Rutherford, K., Salzberg, S.L., Craig, A., James, K., Eisen, J.A., Shallow, S.J., Suh, B., Peterson, J., Kyes, S., Chan, M.S., Nene, V., Selengut, J., Haft, D., Anguoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B.
T	I	T	TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum
J	O	N	JOURNAL	Nature 419 (6906), 498-511 (2002)
P	B	M	PUBMED	12368864
R	E	F	REFERENCE	2 (bases 1 to 250531)
A	A	H	AUTHORS	Hyman, R.W., Fung, E., Conway, A., Kurdi, O., Mao, J., Miranda, M., Nakao, B., Rowley, D., Tamaki, T., Wang, F. and Davis, R.W.
T	I	T	TITLE	Direct Submission
J	O	N	JOURNAL	Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
R	E	F	REFERENCE	3 (bases 1 to 250531)
A	A	H	AUTHORS	Hyman, R.W., Fung, E., Conway, A., Kurdi, O., Mao, J., Miranda, M., Nakao, B., Rowley, D., Tamaki, T., Wang, F. and Davis, R.W.
T	I	T	TITLE	Direct Submission
J	O	N	JOURNAL	Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
F	E	A	FEATURES	Location/Qualifiers
S	O	R	source	1..250531
			/organism="Plasmodium falciparum 3D7"	
			/mol type="genomic DNA"	
			/isolate="3D7"	
			/db xref="taxon:36329"	



Db	172576	ATG-----::: :::	172553
Qy	156	TyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAla	175
Db	172552	TATATGAACAAACAGCAATATATTAACAGAAATATAATATCTTTTA-----	172502
Qy	176	ProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLys	195
Db	172501	-----TATTCAAATGAGAAATGTTTAATAAATCAGCAATAAAGAAAAATAAATAA	172448
Qy	196	AlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSer	215
Db	172447	AGAAATATGATATAGAGAAATGATGNAG---AAATATAT-----	172409
Qy	216	LeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe-----	231
Db	172408	-----GTTAATAAAGCTGTAATAATAATATTTTAAATATCAATAT	172370
Qy	232	IleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyr	251
Db	172369	ATATATGAAGAATTAGATAGATTTTAATCAC-----ATATCTCTTCAAGT-----	172322
Qy	252	GlyLysPheSer-----LysAsnIleAsnPheLysLysIleLysIleLys	266
Db	172321	AAAAAGAGAACCAATTACAAAATGTGAAATGAATATATAAATAAACTTCTTATTTA	172262
Qy	267	GluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHis	286
Db	172561	AAAAATTTCACACGATATGAACAAAATATACCGGCTATTATTTTTTTATAAAGCAT	172202
Qy	287	-----TyrPheLysGlyLys	291
Db	172201	TTACTACATGATTATTTGTAATAAAG 172175	
RESULT 99			
LOCUS	CEY51H4A	269619 bp	DNA linear INV 19-JAN-2006
DEFINITION	Caenorhabditis elegans YAC Y51H4A, complete sequence.		
ACCESSION	AL132952		
VERSION	AL132952.1	GI:6434491	
KEYWORDS	HTG.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	1	(bases 1 to 269619)	
AUTHORS	C. elegans Sequencing Consortium		
CONSTRM	Genome sequence of the nematode C. elegans: a platform for		
TITLE	investigating biology		
JOURNAL	Science	282 (5396), 2012-2018 (1998)	
PUBMED	9851916		
REFERENCE	2	(bases 1 to 269619)	
AUTHORS	Sulston, J.E.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-OCT-1999)		
COMMENT	<p>Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: worm@sanger.ac.uk</p> <p>On May 14, 2001 this sequence version replaced gi:5730145.</p> <p>Coding sequences below are predicted from computer analysis, using predictions from Genefinder (P. Green, U. Washington), and other available information.</p> <p>Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.</p> <p>IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.</p>		

FEATURES	source
1. .269619	Location/Qualifiers
/organism="Caenorhabditis elegans"	
/mol_type="genomic DNA"	
/strain="Bristol N2"	
/db_xref="taxon:6239"	
/chromosome="IV"	
/clone="Y51H4A"	
complement(join(2293. .2346,2393. .2610,3430. .3694, 5250. .5357))	gene
/locus_tag="Y51H4A.1"	
complement(join(2293. .2346,2393. .2610,3430. .3694, 5250. .5357))	CDS
/locus_tag="Y51H4A.1"	
/standard_name="Y51H4A.1"	
/note="contains similarity to Saccharomyces cerevisiae Suppressor of mutant AC40 subunit of RNA polymerase I and III (high serine); SGD:YKR092C"	
/codon_start=1	
/product="Hypothetical protein Y51H4A.1"	
/protein_id="CAB61134.2"	
/db_xref="GI:34555981"	
/db_xref="UniProtKB/TrEMBL:O9NAE5"	
/db_xref=" WormBase:WBGene00013098"	
/translation="MLLVIVLILFILLFFYFGKLLIFGKQKIVENSIEKINSNGIS LAENSKLEKSSQNDPENAEETEMVIEHTAAPVENLEEEIQIKVPEVEDTEKT PAGESPKTSSCGISSEILSEQISIEFSISPPSPFLDFATSYSPALSNFTPTNIV PRMMEPEPSLHKCIKKRSDIRDLRLRIQSAKRHRFRSEHLQQVQHDHIDY"	
complement(join(12157. .12228,13962. .14312,16459. .16593, 17111. .17131))	gene
/gene="rho-1"	
/locus_tag="Y51H4A.3"	
complement(join(12157. .12228,13962. .14312,16459. .16593, 17111. .17131))	CDS
/gene="rho-1"	
/locus_tag="Y51H4A.3"	
/standard_name="Y51H4A.3"	
/note="C. elegans RHO-1 protein; contains similarity to Pfam domain PF00071 (Ras family (contains ATP/GTP binding p-loop))"	
/codon_start=1	
/product="Hypothetical protein Y51H4A.3"	
/protein_id="CAB63379.1"	
/db_xref="GI:6580310"	
/db_xref="GOA:Q22038"	
/db_xref="HSSP:1A2B"	
/db_xref="InterPro:IPR001806"	
/db_xref="InterPro:IPR003577"	
/db_xref="InterPro:IPR003578"	
/db_xref="InterPro:IPR003579"	
/db_xref="InterPro:IPR005225"	
/db_xref="UniProtKB/Swiss-Prot:Q22038"	
/db_xref=" WormBase:WBGene00004357"	
/translation="MAAIRKLLIVDGGACGKTCLLIVFSKQDFPDVYPTVENIVA VIEVDGKQVELAMDVAGQEDYDRRLPSYPTDVLKCFSDISPSLNIPEKWTPE VRHFCPNVPIILVGNKRDRLSDPQTVRLAKMKQBPVKPEQGRALAEQAGAFLECS AKTKDGIREFVEKATQAALQOKKKKKSCMIL"	
complement(join(21593. .21756,22186. .22384,22466. .22675))	gene
/locus_tag="Y51H4A.2"	
complement(join(21593. .21756,22186. .22384,22466. .22675))	CDS
/locus_tag="Y51H4A.2"	
/standard_name="Y51H4A.2"	
/codon_start=1	
/product="Hypothetical protein Y51H4A.2"	
/protein_id="CAB63380.1"	

For a graphical representation of this sequence and its analysis see: <http://www.wormbase.org/perl/ace/elegans/seq/sequence?name=Y51H4A;class=Sequence>

This sequence is the entire insert of clone Y51H4A. The true left end of clone Y4304A is at 193825 in this sequence. The start of this sequence (1. .115) overlaps with the end of sequence AL713992. The end of this sequence (269514. .269619) overlaps with the start of sequence AL132846.



```

QY      201 ulysThr-----TyrIlySileLysLeuTyrCysLeu----- 211
Db      131814 AAAAACTAATTTTTCAAATATACAAATCCATTTTACTGTACAAATCTAGATTTA 131755

QY      212 ----CysProAsnSerLeuLeuAlaAsnPheIleGluLeu-----AlaProAs 226
Db      131754 ATCAGCATCAAAATTTCTTTT-----TCTTACGAATATTATAAAATGGATGACACG 131701

QY      226 nLeuAsnSerAsnPheIleIleGlnGluLys-----AsnAsnTyrThrLy 241
Db      131700 CATCAACTCGAAATATATTAATTTCTCAACATAAAAAATAATATAAACCAATGGTAAC 131641

QY      241 saspile-----LeuIleProSerSerGluAlaTyrGlyLysPheSerLys-----As 257
Db      131640 AGATTTTTTCAATGAAAGTCTCGATAAAATCCATAAACGTTTTTGAATTTGTAAAA 131581

QY      257 nIleAsnPheLysLysIleLysIle-----LysGluAsnValTyrTyrLysLeuIleLysAs 276
Db      131580 CATTATATTCAGAAATTTAAATTCAGAAAAAAGTGTTTTTTTTTGCAAAAAAGC 131521

QY      276 pLeuLeuArgLeuPro 281
Db      131520 TGTTCATAGCTTCCA 131505

RESULT 100
AE014843
LOCUS      271546 bp DNA linear INV 12-OCT-2005
DEFINITION Plasmodium falciparum 3D7 chromosome 11 section 8 of 8 of the
complete sequence.
ACCESSION AE014843 AE014186
VERSION AE014843.1 GI:23496380
KEYWORDS
SOURCE      Plasmodium falciparum 3D7
ORGANISM      Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 271546)
AUTHORS      Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.I.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE      Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL      Nature 419 (6906), 498-511 (2002)
PUBMED      12368864
REFERENCE 2 (bases 1 to 271546)
AUTHORS      Gardner,M.J.
TITLE      Direct Submission
JOURNAL      Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. .271546
location/Qualifiers
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="11"
repeat_region 128. .152
repeat_region complement(130. .207)
repeat_region /rpt_type=tandem
repeat_region /rpt_unit_seq="(TA)n"
repeat_region 150. .183
repeat_region /rpt_type=tandem
repeat_region complement(313. .373)
repeat_region /rpt_type=tandem
repeat_region /rpt_unit_seq="(A)n"
repeat_region 313. .343
repeat_region /rpt_type=tandem
complement(426. .445)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
complement(485. .504)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(644. .683)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
644. .671
/rpt_type=tandem
complement(786. .836)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
799. .836
/rpt_type=tandem
complement(982. .1091)
/notes="AT rich"
complement(1002. .1024)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
1148. .1175
/rpt_type=tandem
complement(1148. .1175)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
complement(1245. .1286)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(1346. .1412)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
1346. .1373
/rpt_type=tandem
1380. .1446
/rpt_type=tandem
1384. .1433
/rpt_type=tandem
complement(1412. .1446)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
<1534. .>2550
locus_tag="PF11_0457"
<1534. .>2550
locus_tag="PF11_0457"
1534. .2550
locus_tag="PF11_0457"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN36037.1"
/db_xref="GI:23496381"
/translation="MEKEKIYDRQLRWGVAQNRMKSNVLVGLSGINIELCKNLI
LNGNITIIDNNIVDEEDIENIFLNEHDMKEYMSVPIFKELKSNQLINIKSFIGHI
NLDTNIIIEKELAYKNEIIQBEKILNDCSDVIDCYITNTVCICSCEDYLYKLTGK
INEICHEKNIGCFANMCHGKYAFLEPSFGKHIIIESYKIKNDNDKKGNDNSIEVOYC
TLSHFLKVPFNSLDKNTDIIYYVFALILYSDKKMNKNKONKOIDEQFLNLYNKLANK
TNTQNTELCKTYKINFSPSCSIMGVTSQEIIRKFIHQHESIPNFCVDFMQNIVCT
SMIK"
complement(1776. .1830)
/notes="AT rich"
/rpt_type=tandem
2544. .2577
/rpt_type=tandem
complement(2555. .2620)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
2575. .2620
/rpt_type=tandem
complement(2664. .2702)
/notes="AT rich"
/rpt_type=tandem
2720. .2748

```



```

Qy      233  ILeGlnGluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrCly 252
Db      113025  AAAAAATAGATGATTATATATAATAAGATATTTTCATTAGATTTTAAAGAAATCTTACT 113084

Qy      253  LysPheSerLysAsnIle-----AsnPhelLysLysIleLysIleLysGluAsnValTyr 270
Db      113085  AAAAAAGGAAGATATATTACATCCCTTAAATAATGATGCTAGAAAAAATATATAT 113144

Qy      271  TyrLysLeuIleLysAspLeuLeuLeuProSerAspIleLysHisTyr 287
Db      113145  ATAAATTAAAAAAGAAACAAAAAATATATATAACGAATGAAGATATAT 113195

RESULT 101
AF343914      15671 bp      DNA      linear      BCT 26-MAR-2002
LOCUS      Campylobacter jejuni NCTC 11828 LOS biosynthesis cluster, partial
DEFINITION
ACCESSION      AF343914
VERSION      AF343914.1 GI:13123726
KEYWORDS
SOURCE      Campylobacter jejuni
ORGANISM      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
              Campylobacteraceae; Campylobacter.
REFERENCE      1 (bases 1 to 15671)
AUTHORS      Oldfield,N.J., Moran,A.P., Millar,L.A., Prendergast,M.M. and
              Ketley,J.M.
TITLE      Characterization of the Campylobacter jejuni heptosyltransferase II
              gene, waaf, provides genetic evidence that extracellular
              polysaccharide is lipid A core independent
JOURNAL      J. Bacteriol. 184 (8), 2100-2107 (2002)
PUBMED      11914340
REFERENCE      2 (bases 1 to 15671)
AUTHORS      Oldfield,N.J. and Ketley,J.M.
TITLE      Direct Submission
JOURNAL      Submitted (30-JAN-2001) Department of Genetics, University of
              Leicester, University Road, Leicester LE1 7RH, UK
FEATURES
     source
         1..15671
             /organism="Campylobacter jejuni"
             /mol_type="genomic DNA"
             /strain="NCTC 11828"
             /db_xref="taxon:197"
             <1..459
             /gene="gmbA"
             <1..459
             /gene="gmbA"
             /note="phosphoheptose isomerase"
             /codon_start=1
             /transl_table=11
             /product="GmbA"
             /protein_id="AAK12965.1"
             /db_xref="GI:13123744"
             /translation="EC1KGGKILICGGSADAQFAAEISGRYKKEPKALAGIAL
             TTDTLSALSGNDYGFVFSRQVEALGNDVLIIGTSGKSPNVLEAFKAKELNML
             LCILGSGKGGMMKLCNHLWPSDDTPIQEMHILIIHTLQIIDEGF"
             complement (463..1536)
             /codon_start=1
             /transl_table=11
             /product="unknown"
             /protein_id="AAK12964.1"
             /db_xref="GI:13123743"
             /translation="MSPASRNNISFYIALYSRFSWFPHPPYFRKETYMKDMRMQARGILI
             LYKDFSAITGTMLENTSAEANKLGGWNINLSLGGSTPALRAVILDYLFKHKDILK
             NIYLSDIRALNELETPKDNFISLYNDKIDLEKLYLSRFINCAIFSKKCKIGK
             NDLDTLNLENKHFGGIEYKSEKWHDKNFQNEILQANFQPNFDIDISHFNVT
             KYLLGFIKYQKQFYLIIPSYSLRYKLSYGEYKDSVLFSNYAILSLWIQET
             QKYPNVKIYGFDDIDYADNIKNYKYPDPAPHYNTDMNSQLNAIRDNTHTLNTQNI VKYLN
             TMRKIKEFOLTPPEVIKQNQF"
             complement (1587..3086)
             /notes="similar to AlgI"
             /codon_start=1

```

```

/transl_table=11
/product="alginate O-acetylation protein"
/protein_id="AAK12963.1"
/db_xref="GI:13123742"
/translation="MLFSSFTFIAPFLPBLTVGVFYLKALQIYTSKAFILILASLFFY
AYFIDYIFILFTLVNFFLANLILRKENNRAGYTLVLYGIIIFNLCLLAIKRYTDF
LENLFLPSKLAHLDFNIPPLHILPLGISFTFOQIAFLVDCYKKSNIEDLLENKVD
FYDCLFTSPFOLIAGPIVHKEMMPQFSLNKEKALIDWKLIAGKGLFISIGLFFK
KYIADSPASWANAAGFTVWONGFLNIIFEAWATLSYTFQLYDFDSGYCDMAIGLALF
FGIALPVNPSPYKALNIADPWRKWHIILGRFLKDYLIYPLGNGRFGKINLNLFIIV
AFLSGVHGAAGFVIMGILHAIAMSHRIYSFWAKGTFLLNSKIYKILISWITFNFI
NIWIFPSENLOQANILLKMGFVIVELPLKFKAGLFLFKKIGGNDTWGVLVSFI
LCIFQKSNLMQDFKSSFKNSLIWMFLFYIALITLIATPYTEFIYFNF"
complement (3138..4088)
/gene="waaf"
complement (3138..4088)
/gene="waaf"
/note="heptosyltransferase II"
/codon_start=1
/transl_table=11
/product="Waaf"
/protein_id="AAK12962.1"
/db_xref="GI:13123741"
/translation="WKPIHLPTWLGDAVMASPALYAIKEHFKNTOFIYGLSVSTAL
PREPNSKIIITENQRYKQALSRLKELGKIDLSFAFRSAFSSKIILHILKTKQRYFF
DNKHKHEHOVLKYLITFENSLSIKAHFKDLKLPKLFQNPILKNGKILGLNPGA
SFGSAKWDASYFAKVALVNFQSOSHEILIFGAGKAEQELCNELYOILKEONI KVNLCN
KTIITLCONIAFCDFLTINDSGPNHISTYKIVAIRGPTKFTQTSFWQONAKLV
HLDIACMPCMQKTCPLKHHKCMKDLKPEKVBEEKLSTP"
4147..4965
/gene="waav"
4147..4965
/gene="waav"
/codon_start=1
/transl_table=11
/product="Waav"
/protein_id="AAK12948.1"
/db_xref="GI:13123727"
/translation="MPQLSIIIPLENSCDFISRVLOSCINQILKDIEIILIDDKSKDN
SLNVLFEAKDPRIKIFQNEBENLGTFASRNIGLVHSSDFIMFLSDDFLTPDACEI
AFKEMKGFLLCFDAFVRVTKQFYRFKQDEVNFQKFLDPLFSQRHFCWSWAKC
FRKDWILKIFEKIDERLSYGEDVLCVIFYMFCEKIAVFKTCIYHFEFNPNGRYEN
KNOEILNQNODKKSNBI IKRLSREFLDFEPHQKLFVLEREEKSLCMRANKI "
complement (4968..5927)
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK12949.1"
/db_xref="GI:13123728"
/translation="WIDNEITVTAFYDIGRDKIKNFERDNDKLSYFELAGIKNKM
IITYQENIKKILDTTRKKHNEEDKTIITKELQEBDFEQYKKIIDFRNYDQSIINRKN
PNNIECSPMYCYLMYLPFFVCDAIQRLTDENIMWDFGNHNGNFVVDKQFNFY
LQEQNSIDENKINFLSKNDKQTLANIYFSMETLMGGIIIFAKPKWLLFKHHMQKC
IKYFTSGIIDDQIMLWLCARNYKYNINIKVYFWMGDIIPQNI AKLKLINTNQ
IKYKIKKEKLKEDFNKNKINIKINILIKYCYFKFINKNHVL"
complement (5924..6880)
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK12950.1"
/db_xref="GI:13123729"
/translation="MNNFVLSLYFIYSAFFLNKRRRIIKGILLHQEHENIANYLEN
AIYIKYFNKLDIDQIKRTRNINGKKIIWQFWYQGDINAPCIKCKFSVQRTKGYNE
VVLIDKONIKOYLIIPDPFIYQKIDDKFGEKTIITPSDLRLSLNLYGGIWLDAQMF
LSGETQKELQDDFFIPIHRSTKQPDYKNINFNYNFFSWDEKFKVNIYVNGFTLSNKN
NEIMKIMODILINWYENKLVYYPFMFOILPDLKKKLLNLYITNDTIDHLLQVHA
KDKYSDKLWNIDKNKTSIHSUKIFKKIRKHSMDKILFKDTI"
complement (6873..7646)
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK12951.1"
/db_xref="GI:13123730"

```





Db	22033	GTGATCCGAA	---	ACCACCAAGTATGAAGAAATCCAAACGTTTGTAGTTTAAATGAAGCAATT	22085
Qy	149	yrLeuSerGlyIle	-----	-----AspPheTyrGlnAsnGlyS	160
Db	22090	ACTTGGAAACATTC	CCAGTATTTC	CCAGCTTGGGAAACACGAGTACTTCCCATATGCTC	22149
Qy	160	erSerTyrAlaPheaspThrLys	GlnGluAsnLeuLysLeuAlaProAspPheLysA	180	
Db	22150	AATTGTCTCCAATTGATTGATTATATAACAACCTGTATCTGTGGAATGTGTGATTAAATGT	22209	909	
Qy	180	snAsp-	-----	-----ArgSerH	184
Db	22210	CTGATTTGGGATTTCAAAATGGATGGTTGCCCGAGAACCAAGCTCTGAGCAACTTAAGAGTC	22259	59	
Qy	184	isTyrIleGlyHisSer	-----	-----LysAsnThrAspIleLysAlaLeuGluPheLeuGluL	202
Db	22270	ACTATGCTGGCTTTTCTACAGTGACGAATCGTGGATTAAAGTTATTTCATTAATCA	22329	29	
Qy	202	ysThrTyrLysIleLysLeuTyrCysLeu	CysProAsnSerLeuLeuAlaAsnPheIleG	222	
Db	22330	ATTGTATTATACCAAAAG	-----	-----AAATTTATGGTCGTATGTGA	22365
Qy	222	luLeuAlaProAsn	-----	-----LeuAsnSerAsnPheIleI	233
Db	22366	ATTTCCTGCAAAACCGCTGATATATTTCGGCCAAATGGCAATTTCTTATAGATCAACTAATG	22425	25	
Qy	233	leGlnGluLysAsnAsnTyrThrLysAspIleLeu	-----	-----ileProSerSerGluAlaIat	251
Db	22426	AAAGTGAAGCCAGTAACAGCAGCGAGTTTGGATATTAGCCACATCCCATCAGCGAGTCC-	22484	84	
Qy	251	yrGlyLysPheSerLysAsnIleAsnPheLysIleLysIleLysGluAsnValTyrT	271		
Db	22485	-----	-----	-----GATACATTGCCAAATTCAGTCTAAAATATTG	22515
Qy	271	yrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyrPheLysGlyI	290		
Db	22516	CTAAGATAGTTGAAAAGATTTC	-----	-----TCCCTTACACTATCGCAAAATATATCTATGTT	22568
RESULT 103					
CR382400/c					
LOCUS	CR382400	348034 bp	DNA	linear	INV 17-APR-2005
DEFINITION	Plasmodium falciparum chromosome 6, complete sequence; segment 3/5.				
ACCESSION	CR382400	AL844505			
VERSION	CR382400.1	GI:46361120			
KEYWORDS	HTG.				
SOURCE	Plasmodium falciparum 3D7				
ORGANISM	Plasmodium falciparum 3D7				
REFERENCE	1 (bases 1 to 348034)				
AUTHORS	Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,B., D., Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K., Buckee,C.O., Burrows,C., Cherevach,I., Chillingworth,C., Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corton,C., Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J., Feltwell,T., Goble,A., Goodhead,I., Gwilliam,K., Hamlin,N., Hance,Z., Harper,D., Hauser,H., Hornsby,T., Holroyd,S., Horrocks,P., Humphray,S., Jagels,K., James,K.D., Johnson,D., Kerhorns,A., Knights,A., Konfortov,B., Kyes,S.D., Larke,N., Lawson,D., Lennard,N., Line,A., Maddison,M., McLean,J., Mooney,P., Moulé,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A., Rabinowitch,E., Rajandream,M.A., Rutter,S., Rutherford,K.M., Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R., Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,L., Whitehead,S., Woodward,J., Sulston,J.E., Craig,A., Newbold,C. and Barrell,B.G.				
TITLE	Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13				
JOURNAL	Nature 419 (6906), 527-531 (2006)				
PUBLISHED	12368867				
REFERENCE	2 (bases 1 to 348034)				
AUTHORS	Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K., Berriman,M., Pain,A., Hall,N., Atkin,R., Chillingworth,C., Doggett,J., Ormond,D., Sanders,M., Hayes,R., Hall,S., Quail,M. and Barrell,B.G.				

JOURNAL REFERENCE AUTHORS	Unpublished 3 (bases 1 to 348034) Cheruvach, I., Davis, P., Goodhead, I., Stevens, K., Mungall, K., Berriman, M., Pain, A., Hall, N., Atkin, R., Chillingworth, C., Dogggett, J., Ormond, D., Sanders, M., Hayes, R., Hall, S., Quail, M. and Barrell, B.G.
TITLE JOURNAL	Direct Submission Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
REFERENCE AUTHORS	4 (bases 1 to 348034) Cheruvach, I., Davis, P., Goodhead, I., Stevens, K., Mungall, K., Berry, A.E., Berriman, M., RA Pain, A., Hall, N., Atkin, R., Chillingworth, C., Doggett, J., Ormond, D., Sanders, M., Hayes, R., Hall, S., Quail, M. and Barrell, B.G.
TITLE JOURNAL	Direct Submission Submitted (26-MAR-2004) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT	For more information about this sequence or the Malaria Project, see <a href="http://www.sanger.ac.uk/Projects/P_falciparum">http://www.sanger.ac.uk/Projects/P_falciparum</a> .
FEATURES source	Location/Qualifiers 1. 348034 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="6" 1757..5722 /locus_tag="PFF0820w" 1757..5722 /locus_tag="PFF0820w" /old_locus_tag="MAL6P1.257" /codon_start=1 /product="hypothetical protein" /protein_id="CA24985.1" /db_xref="GI:46361121" /db_xref="UniProtKB/TrEMBL:Q6LF47"
gene	/translation="MFSRRKIKNKIKSFDNDDEEQIDKLNATHNYKNKEFIDTH DRIQLKPFQNNQNNATFASQVNNVNNKMEIKKEEEKQVQNGQYGTCTF IQSNDKEKBEKIKKKGNKWCISIERKKNKEEKRVKNWNTSFHYISEDDNYSITLKK KKKSSSVKKEKKNKNDIHNHNNYNSDNIFKENOSSDNETFNAYNLSYNS TQGEHIYKDHMKIKDKNLSYNNHNIKNEIYLSSEHNHYNGDHTSDNEENDKD IYDDTPYVNDTSYKQKNISSEYVNDYSNNDENVNGLSDNSGVSVSKGLKN NQIFNSEDVFIDDKDNYQSININLQEDYDKDNDEKKLIERIKLNILRKKKE INNYYNNNDNLDDDDYDNFLPMKKNLFFKIKKNKVTVEDDLDDVMNYET INEMKNRLIKKRNERTKNLYEERNTENNVDTNLDINNNYEDYDLYDEKINQIV DNKMLVELKQNDFLYDSKKKKKEEERENKRVKDKLCPKEDNINAYIKSNN EPKEEHHVKNKEYSDDDLSUSREKDKELYSLDNNNYNDKNKLLFEKIQNAYKNGV CNLEIIMYEHIOQLFSYKNIQESNEIKELNSNQVEYNYKIEICKGKEIILICR VFLQFLQILNLI CEKYPEVDALNGLYKLESTPHLIYHNKLKLYIKEYEYBKILTFI NDYIFNSKYKYNLKEKKKIKOLOYLDNNIIQNGNDNVTOINEYNLNEICDVKNM NNSFYMFDGSSNYSNTDLSSTSLLENDLNGDNKNFIORKENKQILKSDALKSDA LKSDALKNDMSKKNLKDQKYFQTLKMDLKKKFLFISINNI FKNVNNYFFNFK VIRYFGLLLKYNHYFNKKSQYFDVLFVFFVCellyWDPLOYFMSKKKKKRLN YFDEDINSNNKSNKNKNLNLDSNMNDNDKHNEDDNKKFFSLDSESPQYD NDEKNIYNTNERTIETDMDSGSSSSNFSLSHTSSCNENNKNLFLDSSENVK ICSNI FKNKLKKKKYKLRKTFPHNPSIKSFWKYKFMDELILYDNI EEEKDILKLY EHIFNKLYELINWNLKQSLKQSHLSRI INDFILYNDNINELBDKIKDKIQTITTF LQCFNTISQKKKNI FLMRCLKFLKSMKNIIQLCLNBEELYDIKVNFNVYLPNCDDG K"K
gene	complement (6540..7217) /locus_tag="PFF0825c" complement (6540..7217) /locus_tag="PFF0825c" /old_locus_tag="MAL6P1.256" /codon_start=1 /product="mitochondrial import receptor subunit tom40" /protein_id="CA24986.1" /db_xref="GI:46361122" /db_xref="GOA:Q6LF46" /db_xref="InterPro:IPR001925"
CDS	
gene	/db_xref="UniProtKB/TrEMBL:Q6LF46" /translation="MISRIINDGSVNGRFGCKINNNIDCKLNFTYAKNTRNNYEMS LEVNKPLTYNFKSIWQGAIENTSYTOLLTKKIQAGVDLTYIASNCASIGSGFLGRYN KHNVLTMQIVRQPNFKSPFEMLNQHLKYIQYAKKISURLSLGTELEITPQKESAM RQWQSYFPHAKVQSGISVFTQDYSFGVSGYIDYLANDYKFGFMHISPSQ EQPQATAA" Join(12564..15172,15385..15927,16064..16232,16353..16553, 16658..16699,16811..16869,16994..17081) /locus_tag="PFF0830w" Join(12564..15172,15385..15927,16064..16232,16353..16553, 16658..16699,16811..16869,16994..17081) /locus_tag="PFF0830w" /old_locus_tag="MAL6P1.255" /note="Strong Pfam hit (2e-22;codon 25-837) to Adaptin N terminal region; putative alpha adaptin-like protein" /codon_start=1 /product="alpha adaptin-like protein, putative" /protein_id="CA24987.1" /db_xref="GI:46361123" /db_xref="InterPro:IPR002553" /db_xref="UniProtKB/TrEMBL:Q6LF45"
CDS	/translation="MIKHSIKGLYCFIDEVNRCKREDEBKQVLQELIKIKKFNKQ ITNYRKKYIWKLIYCHILGYSYLDIIKLISSNFSKCYGTALSLLVDENNE MLNIMVSTIKADIKNNDEIRLAFHAFKINGLIENLYEDILHVTNLYYKPI RKAFICLANIYKKRHDLLRGKNDPEIFKLDQNLNEMFNI FAYNLLYVILIF OKYESMYIYIKKEKKNSLEKNDLAKREELKFDNSSSSGSSNHSRISLHKNE LKRRKSNYIWNQNFYNDLKYVFLIKLEISECTFNEIEIDFYEIKKYINKYIHI LNVLILDLNENKIDEGFYHYSHRVYFPLIKLOMIQOLYDVHNSLQTTININDIMTK IISKPYRQLQGRNENPNNNI PNNRHITANFNSTNNNNNNNNNNNNNNNN KSSNNNSNIFKKSFKSNISYSHSKKNDILRNEKSALIYEIGIIECSNLVNFDD KIEDRNDIICLITSFDTKSNINVIILNSLSKLNKINIKYNMETHIHLNLLNS LDTYIKLQTFNIFMNCYNNWKLINFLHLHPFDIPYIQNEVVIKISLAENFSD LSWYIDVFMIEITHKSIQFVFRVLQITGFEVDGKNSKSGKNNNNNNNN NSNSNNNNNNNNNNNNNNNNNNNNNNNNKDKQVQTYAAMCYRSLSDHF TRIELLDLCITYIIIGSGYLIKDKVPMKEQLKILEKYFAPASNTCKVILMAIMKVYF YDNKLNSVKKILNTCINHSLELQTRACFLNLCNNMNNMLNLLKNPPLYNMKV HENFLIKLLETNKAVIDFNEKQNSLKFDKDKDKSGKHKNHNDENDYNDTSYSH NKVSHSSCNSSSDEDEDEDEDEEENEEKSKRSRNLSSSHNSKSSSVSSH KSKISDNDNEKFLKILCLOKANNINDLWFNACLLDKCLFPNNNVSILMKQKYQ NKGILIFYKNISKAALNSSINIEEQINIEQKIVNNEKIDSGMYVHKIILTVS DIFYNPSYFNTYNAANASFSKPLLLTRFKENKMNETHFNKNYKALITQVHNE DIVMGVPKSKVFLNLYNAFNFYILKIGMLICASGYIHFDPDLNMMENKILIAKIRH EVKGQISVSVSSIKHLNLYNLKIFEIYLKNN" Join(19555..19659,19825..20227,20451..21203,21343..21422) /locus_tag="PFF0835w" Join(19555..19659,19825..20227,20451..21203,21343..21422) /locus_tag="PFF0835w" /old_locus_tag="MAL6P1.254" /note="Although it is moderately similar to yeast nuclear segregation protein bfr1, the evidence is not strong enough to call this protein a bfr1 homologue" /codon_start=1 /product="hypothetical protein" /protein_id="CA24988.1" /db_xref="GI:46361124" /db_xref="UniProtKB/TrEMBL:Q6LF44"
gene	/translation="MTEAENIKIEPDPDAYNEKLSISQSIDIEIKKIDNLOKEIKV AKQGEYENKKKDIKIDCDSDIDKLENERRSLDDIEKKQKHKELKVAQNMK KQIGPENEDIEKKIREIKMTQSTISIKKEKLLIQALNKNPILLSYSKIE ASKYDDEITVPLKSWDSIREKINKRNEKMKRELQNSYQKNNKLNELNLR LEKDKMHHYFMERSITVMEEEKQOYKSFKNLQAKQCKIKEDRKNKLEKOS NYSKLEDIDVLYREELATENLAYLKKLEEFKLEESKQQAQKADREKNGEVTEN KDDNKNNNNKKKKQKQKTFKIDMNLICLYFYTAGINFPVSDICQKLHE KKDMEQKRDSEKVNVTREDLTNKIKIDEKLTFTFKSGEHNKGTKANKVKA" Join(23040..23150,23307..23486,23689..23817) /locus_tag="PFF0840w" Join(23040..23150,23307..23486,23689..23817) /locus_tag="PFF0840w" /old_locus_tag="MAL6P1.253" /codon_start=1 /product="hypothetical protein, conserved" /protein_id="CA24989.1" /db_xref="GI:46361125"
CDS	

```
/db_xref="UniProtKB/TrEMBL:Q6LF43"
/translations="MDSYKRLRTVELOKEKVTSQKNAPSSIERKLGNIYAPKDFDED
STIVCNFPKTKTISCKNEMQWIKVKEKIPSPQENYLVVFCNPFAXALEI
PLYENKLUFTRAVKEKRETLQWINDMITNNFFS"
complement(join(25622..26665,27855..33995))
/gene="VAR"
/locus_tag="PF0845c"
complement(join(25622..26665,27855..33995))
/gene="VAR"
/locus_tag="PF0845c"
/old_locus_tag="MALGP1.252"
/note="pfEMP1"
/codon_start=1
/product="erythrocyte membrane protein 1 (PFEMP1)"
/protein_id="CAG24990.1"
/db_xref="GI:46361126"
/db_xref="GOA:Q6LF42"
/db_xref="InterPro:IPR004258"
/db_xref="InterPro:IPR006209"
/db_xref="UniProtKB/TrEMBL:Q6LF42"
/translation="MGPAPTATTTSSAKDLLEDIGESVQKEAKKQALGRSESVLHG

Alignment Scores:
Pred. No.: 0.000818 Length: 348034
Score: 128.00 Matches: 78
Percent Similarity: 34.7% Conservative: 55
Best Local Similarity: 20.4% Mismatches: 96
Query Match: 8.2% Indels: 154
DB: 13 Gaps: 19

US-10-734-719-9 (1-291) x CR382400 (1-348034)
Qy 2 LysLysVallelleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
Db 151935 AAAATTTTATACATCTAGGAATTTATTATTATTTGGAACGTATAGAATATATAATTTA 151876
Qy 22 ProAsn-----AspPheAspValPheArgCysAsn 31
Db 151875 ATGGATATATGTAACATATTTGATTTACTATATATATATATATATATATATATATATAT 151816
Qy 32 GlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysAlaValPheTyrThr 51
Db 151815 AATATTTTGTCTCAAGATAAATATCATCTA----- 151786
Qy 52 ProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyr 71
Db 151785 -----ATTTTTAAAGAAATAAAAAATTTA-----ATTATTAAGAAAGTCTTAC 151741
Qy 72 GluThrGluLeuIleMetCysSerAsnTyrAsnGlnAla-----His 85
Db 151740 AAGTCATATGCGATATGCTGTATATCTATACGAGAGCGAAATAATGAATATCTACACAT 151681
Qy 86 LeuGluAsnGlu----- 89
Db 151680 TTCAATAATCATGTATTATATATATATAAATAATTTTGAAAAATAAATATATCAAAA 151621
Qy 90 -----AsnPheValLys----- 93
Db 151620 ATATATAAAGGGAGAATATTGTTATATTTGTTAATATTTGCAAAATTTTATGAAAGACCAG 151561
Qy 94 -----ThrPheTyrAspTyrPheProAlaHisLeuGlyTyrAspPhePhe 109
Db 151560 AAAATAAGTGCTTTGTTTATTTTATATAT----- 151531
Qy 110 LysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArg 129
Db 151530 -----TTAAAAAATTTTCA-----TATAAATTTTAAAAACGTTAAATTAATAAT 151483
Qy 130 IleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyr 149
Db 151482 GAAATTAATAATAATAGGACATGT-----CAAGATATATTC 151447
Qy 150 LeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGlu 169
```

```
Db 151446 ATAAATAATATCGATTTTGTATGATATATCTTAAAAACATTACT----- 151405
Qy 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSer 189
Db 151404 AATATTTTGAATAA-----TATGAAAGAATAATATTTTTCATGTTATC 151360
Qy 190 LysAsnThrAspIleLys-----AlaLeuGluPheLeu 200
Db 151359 AAAAAATACGATATATAACACAATCTTTAAATTTATGTCATGGAATACATACATCTTA 151300
Qy 201 GlulysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaIle 219
Db 151299 GAAAAC-----ATATGTTTATAATAAATTTTAAAAAATGAC 151264
Qy 220 -----PheIleGluLeuAlaProAsnLeuAsnSer 229
Db 151263 GATAACTATAATGCAATAGTTAAAAATATAATTAATTTTGTTCATGTAAATTC 151204
Qy 230 AsnPheIle-----IleGlnGluLysAsn----- 237
Db 151203 TTCITTTGCTTATTTAAGAAAAATTGAAGAAATCCATAAATAACATAAATCAACAAAAGTA 151144
Qy 238 -----AsnTyrThrLysAspIleLeuIleProSerSerGlu----- 249
Db 151143 TATTATTTTAATATCTTAAGAATTTTAAAAAGGATAATTTTAAATATCATAAAAAAGAT 151084
Qy 250 -----AlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIle 265
Db 151083 CTGTTTACTATATATATGATAAG---AACAGGAATATAAATCATAATATTCATAAAAT 151027
Qy 266 LysGluAsnValTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIleLys 285
Db 151026 ---GAAAAATATATAATAATAATTAATAATAATTAATAAGACATCAAGAGAAAATGTTT 150970
Qy 286 HistyrPhe 288
Db 150969 AATATTTTC 150961

RESULT 104
LOCUS CS039302/c 2500 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 112 from Patent WO2005019477.
ACCESSION CS039302
VERSION CS039302.1 GI:61847229
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Lewin,J., Berlin,K., Hildmann,T., Olek,A., Beck,S. and Novik,K.
TITLE Methods and compositions for differentiating tissues or cell types
using epigenetic markers
JOURNAL Patent: WO 2005019477-A 112 03-MAR-2005;
Epigenomics AG (DE)
FEATURES
source
Location/Qualifiers
1..2500
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Alignment Scores:
Pred. No.: 8,06e-06 Length: 2500
Score: 127.50 Matches: 64
Percent Similarity: 39.4% Conservative: 51
Best Local Similarity: 21.9% Mismatches: 118
Query Match: 8.2% Indels: 60
DB: 2 Gaps: 15

US-10-734-719-9 (1-291) x CS039302 (1-2500)
```

```
Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20
   ::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1070 CTPAAATAACAATATATATACAAAACAACACTCTAACAAAATTACAATATAAACAA 1011
Qy 21 LeuProAsnAspPhe-----AspValPheArgCysAsnGlnPheTyrPheGluAsp 37
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1010 TACCCAAATACTACAACACACACATATACCAA--AACAAATATATACATCTATTAA 954
Qy 38 LysTyrTyrLeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGlu 57
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 953 AAAAACAATACATAATAAATACTAACACATCCACCTACCCCTCCCACTACACAAAAA 894
Qy 58 GlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMet 77
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 893 AACCACTATAAAAAAACAACACTACCTATACACCAACCAACCAACCACTACCTATACC 834
Qy 78 CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnGlnPheValLysThrPheTyrAsp 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 833 ACCACATATACCAACCACTAACACAAAACATAAAACATATATACAAAATAATACCTA 774
Qy 98 TyrPheProAspAlaHisLeuGlyTyrAsp-----PhePheLysGlnLeuLys---Glu 114
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 773 TAC-----CACATAACTATACATAAATAAATAAATAAATAAATAAATAAATAAATA 726
Qy 115 PheAsnAlaTyr-----PheLysPheHisGluIleTyrPheAsnGlnArgIleThr 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 725 ACCAACACCTACCTAAAACACTACAAACCCCAAAACACTATACCAACCAACCAATAT--- 669
Qy 132 SerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSer 151
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 668 -----TTCCAAAC---CCTAACTATACCATCCAAACATACCCCAAAAAAAC--- 612
Qy 152 GlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeu 171
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 AAC-----TTCCAAAC---CCTAACTATACCATCCAAACATACCCCAAAAAAAC--- 612
Qy 172 LeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsn 191
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 611 -----TACCATAACCAAAAAAATAATATCAC-----AATAAC 582
Qy 192 ThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeu 211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 581 AACTATATCTACAAAATAATATATACTAACAAAACACTACAAAATAAATAAACAACCA 522
Qy 212 CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe 231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 TTCCCAAT---ATACTCTAAACCTATAATCTATAAACAACCAACCAAAAAAACCTA 465
Qy 232 IleIleGln-----GluLysAsnAsnTyrThrLysAspIleLeuIleProSerSerGlu 249
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 464 AACCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 414
Qy 250 AlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysIle-----LysGlu 267
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 -----ACCAAAACACAAAACCAACCAACCAAAATAAATAAATAAATAAATAA 369
Qy 268 AsnValTyrTyrLysLeuIleLysAspLeuLeuArg 279
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 AACTA-ACCTTCAAATCTCTAAAAACAAATTAATAA 334
```

## RESULT 105

```
AE014831
LOCUS AE014831 250195 bp DNA linear INV 12-OCT-2005
DEFINITION Plasmodium falciparum 3D7 chromosome 10 section 3 of 7 of the
complete sequence.
ACCESSION AE014831 AE014185
VERSION AE014831.1 GI:23494991
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 250195)
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
```

Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Angiuoli,S., Berteau,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrell,B.

Genome sequence of the human malaria parasite Plasmodium falciparum Nature 419 (6906), 498-511 (2002)

12368864

2 (bases 1 to 250195)

Direct Submission

Gardner,M.J.

Submitted (13-SEP-2002) The Institute for Genomic Research, 9712

Medical Center Dr., Rockville, MD 20850, USA

Location/Qualifiers

1. .250195

/organism="Plasmodium falciparum 3D7"

/mol\_type="genomic DNA"

/isolates="3D7"

/db\_xref="taxon:36329"

/chromosome="10"

104. .143

/rpt\_type=tandem

<222..>.2233

/locus\_tag="PF10\_0127"

Join(<222..300,491..533,704..759,912..>.2233)

/locus\_tag="PF10\_0127"

Join(222..300,491..533,704..759,912..>.2233)

/locus\_tag="PF10\_0127"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="AAAN35325.1"

/db\_xref="GI:23494992"

/translation="MSILSLFHICSPFLRNCLEKEIYENLCOTPVITLCNHWNCPI  
ICKQIHKNNLKKITGKQKKEIEIRIKCHLCKEIKIKDYKTHINCEYKCKNYIL  
GCEYDKKSKISHEQSCHEHLSCSSCNLFYFKRIFMLTKELKQHLNMFSTYY  
SFYVLLMNTNFRNLNLIKKHPHTNNNNNNYISRKILKNLQSLDHLCLPVS  
IDTTTSSHTNSLKNKVYIEHHIPTNVSRNSTINFLQNQSPSSHHNPLNEL  
SREIYEHNNHNNYNNDDNDNNSNRLSSYMEYNNISVNPLENISIESREKRKK  
KKTIKKYNFNDLNSKPRNNALKMDQLSGVIENDNSNCEFLNITLPRKKNFND  
KNSKDIITIEELDFDNDLSNIYIDNREFFLCDKSCFTNTIKLQELERSLVLVY  
FCCVGMPPVMTLGCISYVVTGKVFAPFSYIVTITIKLSHKFFLKIFNI"

853. .891

/rpt\_type=tandem

2226..2298

/rpt\_type=tandem

2322..2383

/rpt\_type=tandem

2521..2549

/rpt\_type=tandem

complement(<2646..>.6404)

/locus\_tag="PF10\_0128"

complement(Join(<2646..>.6404,4073,4170..>.6404))

/locus\_tag="PF10\_0128"

complement(Join(2646..>.6404,4073,4170..>.6404))

/locus\_tag="PF10\_0128"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="AAAN35326.1"

/db\_xref="GI:23494993"

/translation="MSIISNIEIDMKFFDFEGDKSVYFCKEYAEFLCLDNLISFK  
IDENKSYSSLYDILYPTCILESSVIDDIKENYOKYLNKRCISLFSFIRYVYCD  
NNVFLSTDDNNITHYLVYNNNNIKGNDYDDDIKNKGGVNNNRDILYVERKEDIEG  
GNGNAYSYSNWIILEKKKIWKSENIWVSCFYKDYDDDDYKLIYGVNGLNLYN  
INTYKIKSYKIHNTRITNLKFKNFIISSDVTKNIYIYDIIKREITICEDHMGII  
DLLFLFVNGKNENKAQNVKGINEEDDYISNDESNDSIDIGIITNNNNNN  
NNLKNNDHYNDEKKLHEKNEBKIGFISGNDKIMNIIDILINSNEEDMNNLISN  
NNLSKKKNSKNKNNKNNLLKFNPIKQFVLSDDINHAIIIPSIIFONENDEKKKK  
KKFIITDENIKWILLTHDNDGULLFYNPILKGDVIYKKNENINTINSISKFLKYF  
LYIFREDSLLYIRYFKLKNYICKLGIFILFNTNSEQLTCDQKYSIHNKHAD

EDDEBEEDMLYKTSNDEKKEKCVILLGDNIIRLRFKNLISORLLICHEDIVSC  
IKLNEKLLFSAANDKNIIFWISLRNYCYLYLKONLYPINAIDVNIKHPESITLVS  
CESSLRLLKCTIDEKLLKKNKHNKCKRLEDFDNKESQSTLTITYPHKVINDIAI  
SPNKLVAUCSTKQRTFETQNLKLIHLEGGHKSQVNYFFSKNDNLLYSNTECVR  
IWSLNFECLEKSTQSLDFNITRVLILNDSMGINGSNGNTSIIINKNCEKLLTNYHK  
DKTFCIQSYEDENKNNNIIIGASINGELIFFDKIDSDKIITEHFVEKRNII FYENCLNL  
LREKKINESLLI CLKLDKPKFTIIVENLYNYTFSILNLIKYOHEYGIGKOLINKOK  
QIUNNDIIKEQNFQSMKNLNNTLLESSMDTOTSINNEANONI VNTTGDYNNMDKN  
NFYNI FENSPLESYIFKELSYMQKGLSEAMDDQNKETE INKSNADNKKDEKND  
PSVKHTNNSDNDSDNYDKEDDNFVLFKKMKKCKSEHSYFLLNKLMEFITFFV  
NHRCAVSNFLNLSIIKIYIDHEDICKINNYQYFPEVYNSYIPRHKNYRVKSLOKSKCF  
ELININYKGDIKMMS"  
4097. 4147  
/rpt\_type=tandem  
repeat\_region  
5368. 5428  
/rpt\_type=tandem  
repeat\_region  
5382. 5428  
/rpt\_type=tandem  
repeat\_region  
6439. 6472  
/rpt\_type=tandem  
repeat\_region  
6442. 6492  
/rpt\_type=tandem  
repeat\_region  
6635. 6663  
/rpt\_type=tandem  
repeat\_region  
6827. 6880  
/rpt\_type=tandem  
repeat\_region  
6972. 7024  
/rpt\_type=tandem  
repeat\_region  
7160. 7184  
/rpt\_type=tandem  
repeat\_region  
7629. 7670  
/rpt\_type=tandem  
repeat\_region  
7632. 7693  
/rpt\_type=tandem  
repeat\_region  
8162. 8191  
/rpt\_type=tandem  
repeat\_region  
8699. 8726  
/rpt\_type=tandem  
repeat\_region  
9127. 9151  
/rpt\_type=tandem  
repeat\_region  
9320. 9358  
/rpt\_type=tandem  
repeat\_region  
9433. 9465  
/rpt\_type=tandem  
repeat\_region  
9768. 9800  
/rpt\_type=tandem  
repeat\_region  
9887. 9912  
/rpt\_type=tandem  
repeat\_region  
10173. 10206  
/rpt\_type=tandem  
repeat\_region  
10209. 10247  
/rpt\_type=tandem  
repeat\_region  
10634. 10664  
/rpt\_type=tandem  
repeat\_region  
11096. 114290  
/locus\_tag="PF10\_0129"  
/locus\_tag="PF10\_0129"  
11096. 14290  
/locus\_tag="PF10\_0129"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AA35327.1"  
/db\_xref="GI:23494994"  
/translation="MNENTMNDNDNTSLVHNLKIRKPRSYDPEYVDDDKTKGRGRK  
KRKRSSRGDINSINGVNDTKENILIKIERTDSNIPDSNENVENKIEIKTECD  
INKCYVSEKTSFNDPIRRYTKKENLEKSMMLFNTNIPLRNNNIKIDRICYLN  
TLNPMRYMGSTAKRFLVEDQOEKYLFLKNNKNLIPEIIKIAMFALLKVLNLDNNIY  
TLPPEKKKEFTISKDMSIKYTKREEYNNNNHNNNNKINDNQNDKTKGDDNNYT  
NNNTVHNPHNNKSNRNNYVAYFYFSKKPNNI ENNEESCDNNNSVKNESNGVKDAS  
KEDETISDIKONCVNSIKSNETNENHTNNYSMTIKIISPACENENISYPYSHYDNSK  
INTSRLEEKDDILNQGVTPPEILDFFVKEKYEKYENVKQYIVTTLNICALNIFKLI  
RGRYITICKYENDFIKMKYFKYIKFYSLKGBEEILMNVNKNYLSKSFYDRNKKKR

RRRSTGIFGFKKRLGKRGSTYCLKEMDALTGKNTIRRRSSSTKKLGRKGLSAVK  
KEEYTNLAVSYTNNNNGVDSNTIGDSSNIIDTNNIMSSNIIMNSSIMDSSNIM  
SSNTDSSNIMDSSNIMDSSNIMDSSNIMDSSNIMDSSNIMDSSNIMDSSNIM  
VDSNNMDCSKDNKIDITHENINIQNDPTNNINNVIKCESSINNVIOSTTELN  
NNGLNDVNTPSYNTIKEDVESNLINKENKSNILNNEELLTNTNDLNKHEEKKD  
BHATLNCI VNNNLDLSNERIKNYCIDLSKINI KICIVVYGMKHPFCSYDQEI SDP  
YFEI IYQVKYDILRRKIERKNLARGANNPSSRRLLIANNNNDNTVKGKNFLL  
RASVSSKDDPKRNSOYKQNEEDDYAMKHEKYNVNFNLSYKTLRKKMICLRH  
VANANTNTNANTNININISTNTNINLVKKEENKIKVKTEIKIKDKCLTPTNES  
LIQSSINSFNSNIISVDFPSINLESCHSQSKA"  
12878. 13173  
/rpt\_type=tandem  
repeat\_region  
12878. 13140  
/rpt\_type=tandem  
repeat\_region  
14016. 14098  
/rpt\_type=tandem  
repeat\_region  
14363. 14410  
/rpt\_type=tandem  
repeat\_region  
14598. 14650  
/rpt\_type=tandem  
repeat\_region  
15351. 15385  
/rpt\_type=tandem  
repeat\_region  
15462. 15522  
/rpt\_type=tandem  
repeat\_region  
15689. 15726  
/rpt\_type=tandem  
repeat\_region  
complement(<15733..>17619)  
/locus\_tag="PF10\_0130"  
complement(<15733..>17619)  
/locus\_tag="PF10\_0130"  
complement(15733..17619)  
/locus\_tag="PF10\_0130"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AAN35328.1"  
/db\_xref="GI:23494995"  
/translation="MKIYRKSTFLLIIFEIFFCVILQNEFVKGSYKGYPLIHL  
NNEYTGSAANFIKNLILNLYIKDVKDFEKTIGDSFLDI EDDTKIERLIKERWL  
YINKINSQMIKSGFIYVHIEKKNFYFLHFLMTNNNGLIILLEDMYINVDE  
YYINKINKLNDQNKKNVSDLLLEKRMFVQSILNNKLOOSIYFIKNNKDIEKIY  
ENYQSKFAYFDLTRNVSIPISKIQNVYKMSKSLYFLNKNNNIHINTFNKQDNO  
TFLTKQKTIICADNFVNIFSDLPSTFTTNSLTITQNLIKLFYDYNKEQVNYN  
ILFFFTNYYVGIQDFLNSINVIFKENIBFVICLDNLNNSDFVIYESNKTPQNVQAFY

Alignment Scores:  
Pred. No.: 0.000689 Length: 250195  
Score: 127.50 Matches: 68  
Percent Similarity: 36.7% Conservative: 47  
Best Local Similarity: 21.7% Mismatches: 115  
Query Match: 8.2% Indels: 83  
DB: 13 Gaps: 15

US-10-734-719-9 (1-291) x AE014831 (1-250195)

Qy 1 MetLysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArg 20  
Db 146311 TTGAGTAAATATTTTGAAGGGTAAT-----TATTCCTTA 146346  
Qy 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyPheGluAspLysTyrTyr 40  
Db 146347 ACAAAAGATGATATAAATATACAAATGTAAGGTATGATATTTGAATCTAAATCAT 146406  
Qy 41 LeuGlyLysLysCysLysAlaValPheTyThrProAsnGlnPhePheGluGlnTyrTyr 60  
Db 146407 TTAAT-----AATACAAATGATCAAGTGAAGATTATTT 146442  
Qy 61 ThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsn 80  
Db 146443 AATATGTTATTTTTCACAAATAAGAAATATTTAATAATATATATATTTGTTAAAT 146502  
Qy 81 -----TyrAsnGlnAlaHisLeuGluAsnGluAsnPhe 91  
Db 146503 AAAAGAATATAGTTTATTTAACTTCTTAATCATTTAAGTGAGAAAGAACCC--- 146559





```
/gene="MSV008"
/codon_start=1
/product="ORF MSV008 leucine rich repeat gene family
protein, similar to Amsacta moorei entomopoxvirus Q3 ORF
SW:P28854"
/db_xref="GI:4049899"
/translation="MNNINLYCYENILKYLDNDKNNLGIYELPKYNCILKVPYY
KYFTYNLEYIKNLCSNCTIDPKFLESINLEILDISYNNKSNLYKCKLSLLKNL
NCNECKIIDPKFLESINLEILDISYNNKSNLYKCKLSLLKNLNCNECKIIDPKFE
SLINLEILDISYNNKSNLYECTPLISLIELKICISCKIIDPKFLESINLEILDISV
YSNIDCKPLVSLKNCFCYIYDFKLETINIKILINISYNNKSNISDCKLPVSL
IKLNCGYIYDFKLEPLINLOKLNISNSYSNICKLPISLIPKSEYCDITNFK
FLRLINLETITVMGRSNTSICKLPISLINLNFIEVDIKNIFKLEINLKSIALS
SRSRNTSQCLPISLVLVDSCHIKNFKLEPLINLKLINISDNGSLNSKCKLPI
SLIEYCSKKNKFNFLQPLINLOKLDISVNDYFSNLFNHPYPPIN"
10282..10476
/gene="MSV007"
10282..10476
/gene="MSV007"
/codon_start=1
/product="ORF MSV007 hypothetical protein"
/protein_id="AAC97609.1"
/db_xref="GI:4049649"
/translation="MYIVNVSKLINRSKNLKFVISOYSDLINLINDIGNLHDFEYEL
ELIFNFKLINGSKLNLSYI"
complement(11374..12138)
/gene="MSV009"
complement(11374..12138)
/gene="MSV009"
/codon_start=1
/product="ORF MSV009 leucine rich repeat gene family
protein, similar to Amsacta moorei entomopoxvirus Q3 ORF
SW:P28854"
/protein_id="AAC97858.1"
/db_xref="GI:4049898"
/translation="MNNLYYKEKLLKLNLDNKLTHKLAERNKLYRYKNSILNIPLE
NHIIYKPKYVRKQCSGKIENNLEKLNLECLTICASKIESKLPISLKKIVFNCK
ITDFKLENLNLEKLIINNKNISIDCKLPISLIPCCNKCEITDFKLELDYNLK
LTYSGNKNISHCINPILKTELECSYCNIDPKFLKLNLYNLKLFITGNNSNISN
CELPKLIKVICNCTIDFNFLKSSNNLRITISN"
complement(12364..14199)
/gene="MSV010"
complement(12364..14199)
/gene="MSV010"
/codon_start=1
/product="ORF MSV010 leucine rich repeat gene family
protein, similar to Amsacta moorei entomopoxvirus Q3 ORF
SW:P28854"
/protein_id="AAC97857.1"
/db_xref="GI:4049897"
/translation="MNNINLYCYENILKYKLDSDHNLASICELFQYYNQINICVNH
YKNYKPKYIRNFSRDRGLTIDTFLEELNLEILDISNNEKLNISCKLPNKI
KKFICVCDIDPKFLEPLINLEILDISYNNKSNLYKCKLSKQKKEICENITDFT
FLKLNKLKULDISENISISKCKLPKTIKFKICDRCIDTDFLEBLINLEILDIS
NLIISNYSKLSKTIKKFICARCAITDFLEBLINLEILDISVNNHKLINSECELP
SLIKLCNCKIKNNLTKLILNINLISFNKIDPKFLENLNLNLEILDISENKNS
ISCKLKTIKKFTCSKAITDFKLEHLNLEILDISYNEANISEPKLSKELN
CSDCYTKDFLERLINLOKLNISYTFSSNISCELTSTLVELNCSCTCKNNFTFIE
KLYNLKILDINNPNNSISECKLTALIELDCTICNIDFKLEPLINLOKLNICGNH
DSNISICKLPUSLELMCMYIISDFKLEPKLSKILINICKNGSSNIYINLSPLE
TIFYAKHDIDFEYINNSLKLIVDYFSDSDSDSDSDSDSDSDSDSDSDSDSDSD
14293..15810
/gene="MSV011"
14293..15810
/gene="MSV011"
/codon_start=1
/product="ORF MSV011 leucine rich repeat gene family
protein, similar to Amsacta moorei entomopoxvirus Q3 ORF
SW:P28854"
/protein_id="AAC97610.1"
/db_xref="GI:4049650"
/translation="MNNINLYCYENLLKYLDKNNKHLAKVSKIFQYYNQINPIHY
```

```
SYFPHYSKYIKILNCEKNIKDKFLEKLSNLEILINISDNKNSNISKNLPNKIKF
ICVRCDIDFNFLERLFRLQMLDISYNYIKSNISCKLPISLIELYKCTNNNFYFL
NDLPLEILDISCNKKIININIRLSKLIKNCSNNNISIDIKFLEHLSNLEILNISY
NKISVYNYKLPNNLIELCKYNNIKNFKFIEBLHNLKLNISNNKSNISCKLPOT
LINLIKACDITDFSFLEPLNLEKLNITEENYDNETNLSKCKLPVSLVLSKLPDNU
CNYLFERLINLEILYINDFPFESILIRSLSPISLREINLSSCYIDDIKFLPELINL
EKNLSDNNLNKITECKLPISLIELDCSDITDFKLEPLINLKLINISDSRNLKII
ECKLPISLIELDCSLCINDFKFLEPLTNLKTITISFNTNILECNLPSNLKIINK"
15855..16247
/gene="MSV012"
15855..16247
/gene="MSV012"
/codon_start=1
/product="ORF MSV012 hypothetical protein"
/protein_id="AAC97611.1"
/db_xref="GI:4049651"
/translation="MNNINLYCYEISNICKFFSFKNLKLVILELHLSNLELPGNLH
SDIFEFSLQNIIFNIFHLKDSKNLSIISLIQSNSINDIGNLQLODIFEFLLYDIINF
FKLCNFSKNVKSIIILQDEQSNILIDEGI"
complement(16436..16654)
/gene="MSV013"
complement(16436..16654)
/gene="MSV013"
/codon_start=1
/product="ORF MSV013 leucine rich repeat gene family
protein, similar to Amsacta moorei entomopoxvirus Q3 ORF
```

## Alignment Scores:

Pred. No.:	0.000755	Length:	236120
Score:	127.00	Matches:	75
Percent Similarity:	35.4%	Conservative:	39
Best Local Similarity:	23.3%	Mismatches:	107
Query Match:	8.1%	Indels:	102
DB:	10	Gaps:	18

US-10-734-719-9 (1-291) x AF063866 (1-236120)

QY	14	LysGluIleAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPhe	33
DB	7865	AAAGATGTGGAGTTAATACATTTTCAATTTGAAATTTTGT---GAAATA	7815
QY	34	TyrPheGluAspLysTyrTyr-----AATTGTAATTTTATTAATAATCTTAT	40
DB	7814	TGTTACAAAATAAAGTGTATTGTACTGTTACATGTCACAGATAGATAGAT	7755
QY	41	---LeuGlyLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr	59
DB	7754	GAATTAAATGGCAATGTAAA-----AATTGTAATTTTATTAATAATCTTAT	7707
QY	60	TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSer	79
DB	7706	TATGCTTATATTGGAGTTTAAATGTATAAAAAATAGAAATTAATATATTATTGGGA	7647
QY	80	AsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhe	99
DB	7646	AACTATGAT-----TATATAAGTAAATATTGTTTCAAGAATGCAAAATATAATAT	7596
QY	100	ProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeu-----	112
DB	7595	-----CTATTAAAGATTGTTTACAAAAAGATGCATTCAA	7560
QY	113	-----LysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPhe	126
DB	7559	TGTAATATGTTTATTGATTCTGAATTAATAACTATAAATTT---TGTCACATTTGCTATTC	7503
QY	127	AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys	146
DB	7502	-----TATTATTATGTT-----TATTGGTATGGATATAT	7473
QY	147	GluIleTyrLeuSerGlyIleAspPhe-----Tyr	156
DB	7472	TG-GTATATATATCC---ATTGATTTTGTATAACTATAATGGGTAATATTATTAGATATA	7417



```

Qy 157 GlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaPro 176
Db |||
7416 TCTAATAATTATGATATAAATAATTTTCATCATTCACAAATCATTAACAAATTA----- 7363
Qy 177 AspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAla 196
Db ::|
7362 -----AATATTAGTAATTGTGATATAAAAAAT 7336
Qy 197 LeuGluPhe-----LeuGluLysThrTyrLys 205
Db |||
7335 ACTAAATTTTATAAATGATTTAAATTTGAAGAAATTAGATATATCTAACAAATTATAAT 7276
Qy 206 IleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaPro 225
Db |||
7275 ATAATAATTTTCATCAATTA---CCAATACATTAATTGAATTAATAATATTAGTCGTGTTAT 7219
Qy 226 AsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIle 245
Db |||
7218 ATAAAAAATACTAATTTTATAATTAATTAGATAAT-----TTGAAAAAATTAGAGATA 7165
Qy 246 ProSerSerGluAlaTyrGlyLysPheSerLysAsnIle-----AsnPheLysLys 262
Db |||
7164 TCTAAAAAGTTTCATATATATGATATGATATATACATAATATTTTACTTTTAAATTTATTAATA 7105
Qy 263 IleLysIleLysGluAsnVal-----TyrTyrLysLeuIleLysAspLeuLeu 278
Db ::|
7104 TTATTAGATAAAGTAATACTTCTATTACATTTTATTATATAAATAAAATCAAAATTTGTGA 7045
Qy 279 ArgLeu 280
Db ::|
7044 AAAATA 7039

RESULT 108.
LOCUS AE014833 252394 bp DNA linear INV 12-OCT-2005
DEFINITION Plasmodium falciparum 3D7 chromosome 10 section 5 of 7 of the
complete sequence.
ACCESSION AE014833
VERSION AE014833.1 GI:23495098
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 252394)
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.-S., Nene,V., Shalom,S.J., Suh,B., Peterson,J.,
Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
PUBMED 12368864
REFERENCE 2 (bases 1 to 252394)
AUTHORS Gardner,M.J.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
FEATURES
source
1. 252394
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="10"
complement(79..213)
/repeat_region
/repeat_type=tandem
253..301
/repeat_region

```







```
Db 70859 AATAATTTTATATAAAGATTGTG----- 70833
Qy 88 AsnGluAsnPheValLysThrPheTyrAspTyrPheProaspAlaHisLeuGlyTyrAsp 107
Db 70832 -----GATTTTACGGTATACGACCTTTTAAAAATGTAGGAATTCAT 70791
Qy 108 -----PhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeu 124
Db 70790 CCATGTAATTATTTCTTACTATAAAAGACAAAAT-----AATATA 70749
Qy 125 TyrPheAsnGlnArgIle-----ThrSerGlyValTyrMetCysAlaVal 139
Db 70748 TATAATATAGAAAAGTTGAAATTTTAAACCCAGTAGAGTAATTATAA----- 70698
Qy 140 AlaIleAlaLeuGlyTyrLysGluIleTyrLysSerGlyIleAspPheTyrGlnAsnGly 159
Db 70697 -----AAAGAGGAATTTTATTAATAAGATATA----- 70674
Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProaspPheLys 179
Db 70673 ----ATTATGACAAAGAGTTTAAACACAGAGGATTTTGAAGAAGCAGAGATTATAA 70617
Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 70616 AAAGATAGATGGCTTTTAAT-----CATATAAAGAAATTAATATATTAAAAAG 70566
Qy 200 LeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsn 219
Db 70565 ATAGAAAATAAATGCAAAATATCTTTAAAGAAATATGT---AATAGT----- 70521
Qy 220 PheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleLeuGln-----Glu 235
Db 70520 TATCAGGATATAATTTCTGGATGTGATAAGCTTTTATAGTAAATAGAGAATATACAA 70461
Qy 236 LysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer 255
Db 70460 AATAAGCATTAGAAAAGAAATATAATAAACCT-----TGGATA 70422
Qy 256 LysAsnIleAsnPheLysLysIleLysIleLysGluAsn-----ValTyrTyr 271
Db 70421 AAAAGTAGTCATATAFAAAAAGAGTGGTATAGAAAGTAAAGATATGTTTTTGATATCTCT 70362
Qy 272 LysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHisTyrPheLys 289
Db 70361 AACGAAATAGATGATGAAAAAAGTATCCTTAATATAATAAGCATTATATATAA 70308

RESULT 112
AF402141 49704 bp DNA circular PLN 31-JAN-2006
LOCUS Schizophyllum commune mitochondrial DNA, complete genome.
DEFINITION Schizophyllum commune mitochondrial DNA, complete genome.
ACCESSION AF402141 AF275272
VERSION AF402141.1 GI:15077911
KEYWORDS
SOURCE Schizophyllum commune
ORGANISM Schizophyllum commune
REFERENCE 1 (bases 1 to 49704)
AUTHORS Forget,L., Ustinova,J., Wang,Z., Huss,V.A. and Franz Lang,B.
TITLE Hyaloraphidium curvatum: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi
JOURNAL Mol. Biol. Evol. 19 (3), 310-319 (2002)
PUBMED 11861890
REFERENCE 2 (bases 1 to 49704)
AUTHORS Forget,L., Ustinova,J., Wang,Z., Huss,V.A. and Franz Lang,B.
TITLE Hyaloraphidium curvatum: a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi
JOURNAL Mol. Biol. Evol. 19 (3), 310-319 (2002)
PUBMED 11861890
REFERENCE 3 (bases 11104 to 15465)
AUTHORS Lang,F.B.F. and Butterwell,C.
TITLE Direct Submission
```

```
JOURNAL Submitted (06-JUN-2000) Departement de Biochimie, Universite de
REFERENCE Montreal, 2900 Edouard-Montpetit, Montreal, Quebec H3T 1J4, Canada
AUTHORS 4 (bases 1 to 49704)
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2001) Departement de Biochimie, Universite de
REMARK Montreal, 2900 Edouard-Montpetit, Montreal, Quebec H3T 1J4, Canada
COMMENT Sequence update by submitter
FEATURES On Aug 2, 2001 this sequence version replaced gi:9966501.
source
1..49704
/organism="Schizophyllum commune"
/organelle="mitochondrion"
/mol_type="genomic DNA"
/db_xref="taxon:5334"
90..3665
/gene="rnl"
90..3665
/gene="rnl"
/product="large subunit ribosomal RNA"
3981..4054
/gene="trnW(cca)"
3981..4054
/gene="trnW(cca)"
/product="trnW(cca)"
/anticodon=(pos:4016..4018,aa:Trp)
4203..4276
/gene="trnd(guc)"
4203..4276
/gene="trnd(guc)"
/product="trnd(guc)"
/anticodon=(pos:4237..4239,aa:Asp)
4688..4759
/gene="trng(ucc)"
4688..4759
/gene="trng(ucc)"
/product="trng(ucc)"
/anticodon=(pos:4721..4723,aa:Gly)
4965..5035
/gene="trnR(ucu)"
4965..5035
/gene="trnR(ucu)"
/product="trnR(ucu)"
/anticodon=(pos:4997..4999,aa:Arg)
5127..5198
/gene="trng(ucc)"
5127..5198
/gene="trng(ucc)"
/product="trng(ucc)"
/anticodon=(pos:5160..5162,aa:Gly)
6090..7952
/gene="rns"
6090..7952
/gene="rns"
/product="small subunit ribosomal RNA"
8226..8297
/gene="trnV(uac)"
8226..8297
/gene="trnV(uac)"
/product="trnV(uac)"
/anticodon=(pos:8259..8261,aa:Val)
8463..8538
/gene="trnL(uag)"
8463..8538
/gene="trnL(uag)"
/product="trnL(uag)"
/anticodon=(pos:8497..8499,aa:Leu)
8956..9027
/gene="trnR(ucg)"
8956..9027
/gene="trnR(ucg)"
```

Submitted (06-JUN-2000) Departement de Biochimie, Universite de  
Montreal, 2900 Edouard-Montpetit, Montreal, Quebec H3T 1J4, Canada  
4 (bases 1 to 49704)

Direct Submission  
Submitted (24-JUL-2001) Departement de Biochimie, Universite de  
Montreal, 2900 Edouard-Montpetit, Montreal, Quebec H3T 1J4, Canada  
Sequence update by submitter  
On Aug 2, 2001 this sequence version replaced gi:9966501.

Location/Qualifiers  
1..49704

/organism="Schizophyllum commune"  
/organelle="mitochondrion"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:5334"

90..3665  
/gene="rnl"

90..3665  
/gene="rnl"

/product="large subunit ribosomal RNA"

3981..4054  
/gene="trnW(cca)"

3981..4054  
/gene="trnW(cca)"

/product="trnW(cca)"

/anticodon=(pos:4016..4018,aa:Trp)

4203..4276  
/gene="trnd(guc)"

4203..4276  
/gene="trnd(guc)"

/product="trnd(guc)"

/anticodon=(pos:4237..4239,aa:Asp)

4688..4759  
/gene="trng(ucc)"

4688..4759  
/gene="trng(ucc)"

/product="trng(ucc)"

/anticodon=(pos:4721..4723,aa:Gly)

4965..5035  
/gene="trnR(ucu)"

4965..5035  
/gene="trnR(ucu)"

/product="trnR(ucu)"

/anticodon=(pos:4997..4999,aa:Arg)

5127..5198  
/gene="trng(ucc)"

5127..5198  
/gene="trng(ucc)"

/product="trng(ucc)"

/anticodon=(pos:5160..5162,aa:Gly)

6090..7952  
/gene="rns"

6090..7952  
/gene="rns"

/product="small subunit ribosomal RNA"

8226..8297  
/gene="trnV(uac)"

8226..8297  
/gene="trnV(uac)"

/product="trnV(uac)"

/anticodon=(pos:8259..8261,aa:Val)

8463..8538  
/gene="trnL(uag)"

8463..8538  
/gene="trnL(uag)"

/product="trnL(uag)"

/anticodon=(pos:8497..8499,aa:Leu)

8956..9027  
/gene="trnR(ucg)"

8956..9027  
/gene="trnR(ucg)"

/gene="trnR(ucg)"

```
/product="trNA-Arg"
9374..10162
/gene="atp6"
CDS
9374..10162
/gene="atp6"
/EC number="3.6.1.34"
/codon_start=1
/transl_table=4
/product="ATP synthase F0 subunit 6"
/protein_id="AAK83398.1"
/db_xref="GI:15077912"
/translation="MTYNMFLSNDVYLSPLNQFVNTLIGLNI PMNIVNFSFTNLS
LYTCLVVIIGLHFGYGLDFTLIPNNWSITLESFYASISNNVREQISKTSEVIFPI
YSLPPFLITLNI SNVPSPAVTSASVSLGLSMITPIFGVTILALSINHKIFSEFPVP
GSLTALVPLLVLEFVSFAVSUSGLRFLFANLFAHGLAILLSSFLYKLFSSGILV
AILTLIPFALMIGIIGLEAVSFQAFVFTILTCSYKDLAILH"
11104..15465
/gene="rps3"
CDS
11104..15465
/gene="rps3"
/EC number="1.9.3.1"
/codon_start=1
/transl_table=4
/product="cytochrome c oxidase subunit 1"
```

```
/protein_id="AAK83399.1"
/db_xref="GI:15077913"
/translation="MRWLMSTNAKEIGTLTYLIVSLFAGMIGTAFSVLIRLELSPGVQ
FLQGDHQLENVIIITAHAFIMIFEMWMPGLIGGNGYVFLPHCCAIIDMAPRLANNISFW
LIVPSLLLLLSALVNGAGTGWTPYPLSGIGSQSHSGPSVDLAI FSLHSGVSSLLGA
INFITALLNMRNTNGMLHKLPLFVWAI FVTAILLLLSLVLAGGITLMLLDRNPNTSF
FDPAGGDPILYQHLFWFFHPDEVIIIPGFGVISHVVTSGKTFYGLGMVYAMF
SIGILGFLVSHMFPAVMDVDSRAVFTAAVTAVITGTFKIFSWLATLFGGSLRFNT
PLVAMGFLALFTIGLTCGMLSNASLDVAHDTYVVAHFHVLVSMGAVFALFAGFY
FWMPKIVGTYSEFLSNVHFPTLFGVGNLTFFPQHFLGMSGMERRIPDYDAFTGNWL
VSSIGIISVSIVIIIFYIILYDMFAVQGLVSNMPVEPSTFYVDMELFENDEQAGSLE
WSFDPTIAEHAYNMOPVQS"
/gene="trnS(guc)"
complement(18412)..18493
/gene="trnS(guc)"
complement(18412)..18493
/gene="trnS(guc)"
complement(18412)..18493
/product="trNA-Ser"
/anticodon=(pos:18459..18461,aa:Ser)
complement(18609)..18680
/gene="trnE(uuc)"
complement(18609)..18680
/gene="trnE(uuc)"
complement(18609)..18680
/product="trNA-Glu"
/anticodon=(pos:18645..18647,aa:Glu)
complement(18809)..18881
/gene="trnM(cau)f"
complement(18809)..18881
/gene="trnM(cau)f"
complement(18809)..18881
/product="trNA-Met"
/anticodon=(pos:18845..18847,aa:Met)
```

## Alignment Scores:

```
Pred. No.: 0.000302 Length: 49704
Score: 125.00 Matches: 77
Percent Similarity: 35.3% Conservative: 53
Best Local Similarity: 20.9% Mismatches: 98
Query Match: 8.0% Indels: 140
DB: 4 Gaps: 18
```

US-10-734-719-9 (1-291) x AF402141 (1-49704)

```
QY 1 MetLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAsp----- 17
Db 14275 TTACAAAATAATTGTTGTTAATAATAATCTAGTTTAAAATAATTGATAATGGTTGA 14334
QY 18 ---TyrSerArgLeuProAsnAspPheAspValPheArgCys-----AsnGlnPhe 33
Db 14335 AAATATTCA--ATTGATAACAATTATAATTATATTAGATTGATTTAATAATCAAAAT 14391
QY 34 TyrPheGluAspLysTyrTyrLeuGlyLysCysLysAlaValPheTyrThrProAsn 53
Db 14392 TTTGATAAGAATTCAATCTTTTAAATAAGAAATTTTAGCTTTATTCT---ACTAATCT 14448
QY 54 PhePheGluGluTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGluTyrGluThr 73
Db 14449 TTTCTTTGATAAACAATAT-----AATCTCTATTATTGAAAAATCCTTTATTCCTTAAT 14499
QY 74 GluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLys 93
Db 14500 ACTGCTTTA-----CAAAATTAATTCGAAGATAATAATTATGCT--- 14538
QY 94 ThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLys 113
Db 14539 -----ATATTATATGACTACTACTATAAGAAATAGT 14568
QY 114 GluPheAsn-----AlaTyrPheLysPheHisGluIleTyr 125
Db 14569 TATTATAATATGAATCAAAAAGAGTGATTTTCTCTTCAATAAAATATTATTCTTCTTC 14628
QY 126 PheAsnGlnArgIleThrSerGly-----ValTyr----- 135
Db 14629 TTTAATACTAGAAATACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 14688
```

```
Qy 136 -----MetCysAlaValAlaIleAlaLeuGlyTyrLysGluIle 148
Db 14689 GCTCAATCTTTAGACCTATGTTCTGTTAAATAGT-----AAACCTATT 14733
Qy 149 TyrLeuSerGlyIleAsp-----PheTyrGlnAsnGlySerSerTyr 162
Db 14734 TTTGTTATTACTCCGGATAAAGTGACATTTTAAATATTTTATTACTTATTTCCTTAC 14793
Qy 163 AlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsn----- 180
Db 14794 TTACTTTAAATTTTAAAAAAGATATAATTTTAAATTTACTTAAATTTTAAAAATAGTGA 14853
Qy 181 -----AspArgSerHisTyrIleGlyHis 188
Db 14854 TTATTAAAGTAATAAAGCCGTAAGAGAAATTAAAAAGAGACAGATATATTAGAAAA 14913
Qy 189 SerLysAsnThrAspIleLysAlaLeuGluPheLeu-----GluLysThr 203
Db 14914 AGAAAAAATAATAAGAAAAAAGAGATTCTTCTTAATCTAAGAAAAAGAAAAAGAGAA 14973
Qy 204 TyrLysIleLeuTyrCysLeuCysProAsnSerLeu----- 216
Db 14974 TTTAGAAAAAATAATATTATTAGTTTCTTAGAAACAGCTTAAAGTTCTTATTAGAAAAAGA 15033
Qy 216 ----- 216
Db 15034 TTCCAAGTATTATGTCATATTTTAAAGTAATTTTAAAAAGACCTGTTGAATTTAAATTA 15093
Qy 217 -----LeuAlaAsnPheIleGluLeuAla 224
Db 15094 ACTAGATTACATTATCATATACAGATAGTAATATTATTAGTAACTTTCTTAGCTATTATG 15153
Qy 225 ProAsn-----LeuAsnSerAsnPheIleGln 234
Db 15154 ATTAATAAAGCTAAATTAACCTTTATTCTTATTAAACTTTTCAGTAATACAGTATTTAAA 15213
Qy 235 GluLysAsnAsn--TyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLys 253
Db 15214 AGATATAACAATAATTCTGTTAAAGAAAGATGATGATGATGATGATGATGATGATGAT 15255
Qy 254 PheSerLysAsnIleAsnPheLys 261
Db 15256 TTCACTCTCTGTTTAAATATATAGA 15279
```

RESULT 113

PFMAL13\_03/c

WPCOMMENT

Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509

Fragment Name	Begin	End
PFMAL13_00	1	110000
PFMAL13_01	100001	210000
PFMAL13_02	200001	310000
PFMAL13_03	300001	410000
PFMAL13_04	400001	510000
PFMAL13_05	500001	610000
PFMAL13_06	600001	710000
PFMAL13_07	700001	810000
PFMAL13_08	800001	910000
PFMAL13_09	900001	1010000
PFMAL13_10	1000001	1110000
PFMAL13_11	1100001	1210000
PFMAL13_12	1200001	1310000
PFMAL13_13	1300001	1410000
PFMAL13_14	1400001	1510000
PFMAL13_15	1500001	1610000
PFMAL13_16	1600001	1710000
PFMAL13_17	1700001	1810000
PFMAL13_18	1800001	1910000
PFMAL13_19	1900001	2010000
PFMAL13_20	2000001	2110000
PFMAL13_21	2100001	2210000
PFMAL13_22	2200001	2310000
PFMAL13_23	2300001	2410000

```
PFMAL13_24 2400001 2510000
PFMAL13_25 2500001 2610000
PFMAL13_26 2600001 2710000
PFMAL13_27 2700001 2732359
Continuation (4 of 28) of PFMAL13 from base 300001 (AL844509 Plasmodium falciparum 3D7 C
Alignment Scores:
Pred. No.: 0.000651 Length: 110000
Score: 125.00 Matches: 75
Percent Similarity: 36.8% Conservative: 54
Best Local Similarity: 21.4% Mismatches: 94
Query Match: 8.0% Indels: 128
DB: 12 Gaps: 22
US-10-734-719-9 (1-291) x PFMAL13_03 (1-110000)
Qy 23 AsnAspPheAspValPheArgCysAsnGlnPhe---TyrPheGluAspLysTyrTyrLeu 41
Db 57603 AATGATATTGAT--TACACATGTGGTGAATAAATAATATATGAAGATGAAGAGAAAAA 57547
Qy 42 GlyLysLysCys----- 45
Db 57546 AAAAAGAAATGTATTCAAATAATTACAAAAACATATAATTAATCTGTAAAAATACAAATGAACAA 57487
Qy 46 ---LysAlaValPheTyrThrProAsn-----PhePheGlu----- 57
Db 57486 AATAAAATCTTTTATTGTAATCTGACATTTTATTAGATCTATTTTCTTTTACGTTGAA 57427
Qy 58 ---GlnTyrTyrThr-----LeuLysHisLeu-----IleGlnAsn 68
Db 57426 CTGAATACTATAAAGGTATTCTTCTTCTCAAAAGTTTATTGTAGTATTTCCTCAAT 57367
Qy 69 ---GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGlu 87
Db 57366 TTTAAGAAATTCGATATATTTTATTCTACAAAATAATAATATATCATATAATATAAGAA 57307
Qy 88 AsnGluAsnPhe----- 91
Db 57306 AATAAATATTGAATTAGATAAATAAACAATGTTCTCTTATTATATAAAAAATTCATGCC 57247
Qy 92 -----ValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePhe 109
Db 57246 CTGAATACTTACAAATATATGAATATTCTCAAAAGTTGTTTGTAGTACATATGACTTT 57187
Qy 110 LysGlnLeuLysGlu-----Phe 115
Db 57186 ATTAAGATAAAAAAATCTTAGAAGAGAGGTGGAAGAGATCATATAAATAGCATGTAT 57127
Qy 116 AsnAlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyr 135
Db 57126 CATTTTATTTCGA---AATGAAATTTACAAATACTTTTAAAGATACG----- 57082
Qy 136 MetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPhe 155
Db 57081 -----CATACTTATATATCAATAATAAATAATAT 57055
Qy 156 TyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAla 175
Db 57054 AATATGAAT-----GAATTTATTAGAAAAATAAT 57028
Qy 176 ProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLys 195
Db 57027 -----TGAATTGTAAAAAATCAAAACAATCAG 57001
Qy 196 AlaLeuGluPheLeuGluLysThrTyr-----LysIleLysLeuTyrCysLysPro 213
Db 57000 AAAGTACAAATAGTAAATGATACATATATAGTACAGTAATAATATATCTTACA 56941
Qy 214 AsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIle 233
Db 56940 AATTCAAAATA---AATTATATAAGTCCAAATAAAACATTATAAATAGTATATTGTAAT 56884
Qy 234 GlnGluLysAsnAsnTyrThr-----LysAspIleLeu 244
```

```

:::|||||:::|||||:::|||||:::|||||
56883 AAGAGAAAAAATAATATAATGAAGAAACCAATGCGGATTACAAAAAAGAAATCAGT 56824
Qy 245 IleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhelLysLys 264
Db 56823 ATACCAATAGT---TTGTATTAATATATATATCC-----ATAAAATATGAAAAAACA 56773
Qy 265 IleLysGlu-----AsnValTyrTyrLysLeuLysAspLeu---LeuArgLeuPro 281
Db 56772 CTTTITGATATAAATAATATATATATATATATCTTAACAATATATATATATACGCCCA 56713
Qy 282 SerAspIle-----LysHisTyrPhe 288
Db 56712 GTTGATTATTGCTTTACATATTTCATTATTAC 56680

RESULT 114
PFMAL8P1_11/c
WPCOMMENT
Sequence split into 14 fragments LOCUS PFMAL8P1 Accession AL844507
Fragment Name Begin End
PFMAL8P1_00 1 110000
PFMAL8P1_01 100001 210000
PFMAL8P1_02 200001 310000
PFMAL8P1_03 300001 410000
PFMAL8P1_04 400001 510000
PFMAL8P1_05 500001 610000
PFMAL8P1_06 600001 710000
PFMAL8P1_07 700001 810000
PFMAL8P1_08 800001 910000
PFMAL8P1_09 900001 1010000
PFMAL8P1_10 1000001 1110000
PFMAL8P1_11 1100001 1210000
PFMAL8P1_12 1200001 1310000
PFMAL8P1_13 1300001 1325595
Continuation (12 of 14) of PFMAL8P1 from base 1100001 (AL844507 Plasmodium falciparum 3D

Alignment Scores:
Pred. No.: 0.00651 Length: 110000
Score: 125.00 Matches: 61
Percent Similarity: 37.3% Conservative: 46
Best Local Similarity: 21.3% Mismatches: 96
Query Match: 8.0% Indels: 84
DB: 12 Gaps: 13

US-10-734-719-9 (1-291) x PFMAL8P1_11 (1-110000)
Qy 34 TyrPheGluAsp-----LysTyrTyrLeuGlyLysLysCysLysAlaValPhe 49
Db 93085 TATTATAGGATAAGAAGATTATATAATATGAACTTCGAAATGAAATTTTGTAGATTAT 93026
Qy 50 TyrThrProAsnPhelPheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGln 69
Db 93025 AAAAAATTAAGATATTCATATGAACATTTATTATATATGAACATATAATAATAATGTA 92966
Qy 70 GluTyrGluThrGluLeuIleMetCysSer-----AsnTyrAsnGlnAla 84
Db 92965 AAAAAATAAAGAACTAATCTATACATCATTTAGTGGAAGCAATTAATATGAAATTA 92906
Qy 85 HisLeuGluAsnGluAsnPhelValLys-----ThrPheTyrAspTyrPhePro 100
Db 92905 AAAATGAAAAAGATAAATAATTTTAAATATATATATATATATATATATGATTTTATC 92846
Qy 101 AspAlaHisLeuGlyTyrAspPhePhe----- 109
Db 92845 GAAACATATTAGGTATAGATATATATTCGATCGATCAATATATTTTATAACATA 92786
Qy 110 -----Lys 110
Db 92785 AGTTTTTTAGATATATATAAAAAATATTGTTGCTCTTGAATTAATCAATCAATTA 92726
Qy 111 GlnLeuLysGlu-----PheAsnAlaTyrPheLysPhe 121
Db 92725 ACGTTAAAAAGAAATGTTTCACATACACAAAAATAAAATGTTAATAAATATATCTTAT 92666

```

```

Qy 122 HisGlu-----IleTyrPheAsnGlnArgIleThrSerGlyValTyr---MetCys 137
Db 92665 CATGATCATATCAAAATATCTTTTCAAAATCAAGTCGATTCTTAAGTATATATCTACTGT 92606
Qy 138 AlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLysSerGlyLysIleAspPheTyrGln 157
Db 92605 ATCGTAACCGTTCCTATAGGGGTTTTAAAG----- 92576
Qy 158 AsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaProAsp 177
Db 92575 -----GACAGCGTTCGTTTTTTAATTCAAAAAGAAATAAGCGGAAAAAGGGTACCAAT 92522
Qy 178 PheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu 197
Db 92521 ATAAACAGCAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 92471
Qy 198 GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeu 217
Db 92470 ---TATGTCAAAAAATTAATAAGAAAAAATCAAT-----AAAAATAAAATTTATG 92423
Qy 218 AlaAsnPhelIleGluLeuAlaProAsnLeuAsnSerAsnPhe-----Ile 232
Db 92422 GAAAAATAATAATACCAATGTAATAATAATAATAATAATAATAATAATAATAATAATAATA 92363
Qy 233 IleGlnGluLysAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGly 252
Db 92362 ATA-----AAAAACAAT-----GAACAATATTTTAAACAATACAGCTATATC 92315
Qy 253 LysPheSerLysAsnIleAsn 259
Db 92314 AATAATTCAGACACATAAGT 92294

RESULT 115
AC004688 169794 bp DNA linear HTG 12-AUG-2000
LOCUS Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.
ACCESSION AC004688 GI:9797726
VERSION HTG; HTGS PHASE1.
KEYWORDS Plasmodium falciparum (malaria parasite P. falciparum)
SOURCE Plasmodium falciparum
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 169794)
AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
Kurdi, O.B., Conway, A.B. and Davis, R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
2 (bases 1 to 169794)
AUTHORS Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
Direct Submission
TITLE Submitted (16-MAY-1998) Stanford DNA Sequencing and Technology
JOURNAL Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT On Aug 12, 2000 this sequence version replaced gi:8810452.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 43829: contig of 43829 bp in length
* 43830 44029: gap of unknown length
* 44030 87202: contig of 43173 bp in length
* 87203 87402: gap of unknown length
* 87403 94824: contig of 7422 bp in length
* 94825 95024: gap of unknown length
* 95025 169794: contig of 74770 bp in length.
FEATURES
Location/Qualifiers
1..169794
source

```



```
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clones="PPYAC69"
/clones="3D7"
43830..44029
/estimated_length=unknown
87203..87402
/estimated_length=unknown
94825..95024
/estimated_length=unknown
ORIGIN
Alignment Scores:
Pred. No.: 0.00099 Length: 169794
Score: 125.00 Matches: 71
Percent Similarity: 37.1% Conservative: 46
Best Local Similarity: 22.5% Mismatches: 128
Query Match: 8.0% Indels: 70
DB: 12 Gaps: 14
US-10-734-719-9 (1-291) x AC004688 (1-169794)
Qy 14 LysGluLeuAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPhe 33
Db 91983 AACAGTTGATATAAATAAGATG-----GATAAAATTTTATAAATCAAAATACCTTT 91930
Qy 34 TyrPheGluAspLysTyrTyrLeuGlyLysCysLys-----AlaValPheTyrThr 51
Db 91929 TTTATAATAATTTATATTAT--CAITTAAGTGTAAATTTATATGTTGTTCAAAATTT 91873
Qy 52 ProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu-----IleGlnAsnGln 69
Db 91872 ATAAATATGTCAGATGATGATTTTATGACAGATATACATATATCTCAAAATTAAGG 91813
Qy 70 GluTyr-----GluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGlu 87
Db 91812 AGGTATGTAATGAACAGATAGTATATATATACCTTTTAAATATACATACATTAAGT 91753
Qy 88 AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAsp 107
Db 91752 AATGAAGTATGGTA-----TATATATTTTTCGAATTTATATATATCAATATG 91702
Qy 108 PheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn 127
Db 91701 TTT-----AAGGAAATAATTTTCAAAATACGTGTGCATACATATTTACATTTTC 91654
Qy 128 GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGlu 147
Db 91653 CAATTAATAAACAATAATTCACAAAAGTTAAATTTGATTAA-----TATAGTTCA 91600
Qy 148 IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLys 167
Db 91599 AAGTTTGTATTACATATGTTTAAAGAAAAAGTGTACAAAAAATAAATAAATAA 91540
Qy 168 GlnGluAsnLeuLeuLysLeuAlaProAspPheLysAsnAsp----- 181
Db 91539 AATGATCATCATGAGAGGATAGTATATAATAATAATATGATAGTAAATAATAATAT 91480
Qy 182 -----ArgSerHisTyr-----IleGlyHisSer 189
Db 91479 TATAATAGTAATAATAATAGTAATATTTGTGTTCCAAATAATATTTATTTATGACATGAT 91420
Qy 190 LysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyr 209
Db 91419 GAAAATTTTATTTCAATATTTGATTT----- 91390
Qy 210 CysLeuCysProAsnSerLeu-----LeuAlaAsnPheIleGluLeuAlaPro 225
Db 91389 -----AACACATAGATCAAAATTAAGTGCATATTTATTAATTAAGGATGAA 91342
Qy 226 AsnLeuAsnSerAsnPheIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIle 245
Db 91341 ACAATAAATGAGAATTTTATTTTGAAGTACGAAATTAATAATACATTCATCCTTATT 91282
Qy 246 ProSer-----SerGluAlaTyr 251
Db 91281 TCTTCTAAATGTAATATAATAATAACATTCGCGAAAAAAGAAAGAGAGATAT 91222
Qy 252 GlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnValTyrTyr 271
Db 91221 AATAAA-----AAATGAGTAATAATTCATCATGTAGTTTAAATGAAATGTTTATATG 91168
Qy 272 LysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHis 286
Db 91167 GAAGGTAGAAATTTTGAATTTATTTGTTTCATCAATAATAAGAAAT 91123
RESULT 116
AE014847/c 252650 bp DNA linear INV 11-FEB-2003
LOCUS
DEFINITION
Plasmodium falciparum 3D7 chromosome 12, section 4 of 9 of the
complete sequence.
ACCESSION
AE014847 AE014188
VERSION
AE014847.1 GI:23496718
SOURCE
Plasmodium falciparum 3D7
ORGANISM
Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 252650)
Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
Crumholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Frumkin,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
Genome sequence of the human malaria parasite Plasmodium falciparum
Nature 419 (6906), 498-511 (2002)
12368864
2 (bases 1 to 252650)
Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
Direct Submission
Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
3 (bases 1 to 252650)
Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
Direct Submission
Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES
source
1..252650
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"
90..134
/rpt_type=tandem
repeat_region
96..134
/rpt_type=tandem
repeat_region
586..669
/rpt_type=tandem
repeat_region
724..748
/rpt_type=tandem
repeat_region
1220..1250
/rpt_type=tandem
repeat_region
1331..1363
/rpt_type=tandem
repeat_region
1440..1476
/rpt_type=tandem
repeat_region
```



Pred. No.:	0.00145	Length:	252650		
Score:	125.00	Matches:	71		
Percent Similarity:	37.1%	Conservative:	46		
Best Local Similarity:	22.5%	Mismatches:	128		
Query Match:	8.0%	Indels:	70		
DB:	13	Gaps:	14		
US-10-734-719-9 (1-291) x AE014847 (1-252650)					
Qy	14	LysGluLeuAspTyrSerArgLeuProAsnAspPheAspValPheArgCysAsnGlnPhe	33		
Db	210995	AAACAGTTTCATATAAATAGATG-----GATAAAATTTATTTAAATCAATACCTTT	210942		
Qy	34	TyrPheGluAspLysTyrTyrLeuGlyLysCysLys-----AlaValPheTyrThr	51		
Db	210941	TTTATAATAAATTTATATTAT---CATTTAAAGTGAAATTTATTTATGTTGTTTCAAAATTT	210895		
Qy	52	ProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeu-----lleGlnAsnGln	69		
Db	210884	ATAAATATGTCAGATGATGTTTATGACACATATATACATATCTCAAAATTTAAAAAAGG	210825		
Qy	70	GluTyr-----GluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGlu	87		
Db	210824	AGGTATGTAATGAACAAGATGTTATATATATACCTTTTAATATATACATCAATTAAGT	210765		
Qy	88	AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAsp	107		
Db	210764	AATGAAGGTATGTA-----TATATATTTTGTGCAATTTATATATATCAATATG	210714		
Qy	108	PhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPheAsn	127		
Db	210713	TTT-----AAGGAAAAATAATTTTGAATACGTGTCATACATATTTACATTTTC	210666		
Qy	128	GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGlu	147		
Db	210665	CAATTAAATAAACAATAATTCATACAAAGTTTAAATTTGATTTAA-----TATAGTTCA	210612		
Qy	148	IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrTyrAlaPheAspThrLys	167		
Db	210611	AAGTTTGTGTTTCACTATGTTTAAAGAAAAAGTAGTAGCAAAAAAATAAAAAA	210552		
Qy	168	GlnGluAsnLeuLysLeuAlaProAspPheLysAsnAsp-----	181		
Db	210551	AATGATCATCATGAGAGGATGATGATAATAATAATGATAGTAATAATAATATAT	210492		
Qy	182	-----ArgSerHisTyr-----lleGlyHisSer	189		
Db	210491	TATAATAGTAATAATAATAGTAATTTATGTTTCCAATAATATTTATTTATGACATGAT	210432		
Qy	190	LysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyr	209		
Db	210431	GAATAATTTATTTCAATAGTTGTGATTT-----	210402		
Qy	210	CysLeuCysProAsnSerLeu-----LeuAlaAsnPheIleGluLeuAlaPro	225		
Db	210401	-----AACACATTAGATCAAAATTAAGTGGCAATTTATTTAATTAAGGATGAA	210354		
Qy	226	AsnLeuAsnSerAsnPheIlelleGlnGluLysAsnAsnTyrThrLysAspIleLeuIle	245		
Db	210353	ACAATAAATGAGAAATTTTATTTTGAAGTAGCAAAATATATAATACACATTCATCCTTATT	210294		
Qy	246	ProSer-----	251		
Db	210293	TCCTTAAATGTTAATAATAATAATAACATTCGCGGAAAAAATAAGAGAGATAT	210234		
Qy	252	GlyLysPheSerLysAsnIleAsnPheLysIleLysIleLysGluAsnValTyrTyr	271		
Db	210233	AATAAA-----AAATGAGTAATAATTCATCATGTAGTTTAAATGAAATGTTTATATG	210180		
Qy	272	LysLeuIleLysAspLeuLeuArgLeuProSerAspIleLysHis	286		
Db	210179	GAAGGTGAGAAATTTTGAATTTATTTGTTTCATCCAAATAATAAGAAAT	210135		

RESULT 117	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
LOCUS	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
DEFINITION	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
ACCESSION	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
VERSION	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
KEYWORDS	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
SOURCE	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
ORGANISM	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
REFERENCE	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
AUTHORS	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
TITLE	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
JOURNAL	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
PUBMED	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
CONSRMT	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
TITLE	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
JOURNAL	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
FEATURES	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
source	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
gene	Prochlorococcus marinus subsp. pastoris str. CCMP1986	348986 bp	DNA	linear	BCT 17-APR-2005
CDS	Prochlorococcus marinus subsp. pastoris str. CCMP1986	34898			



Db	20442	-----TTTGACTTTTAATAGATGTTTTTATGAATATCTGGAT	20408
Qy	65	uileGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHi	85
Db	20407	TTATGATTTTAAGAATCATATAAATGAATTTTATTAAT	20357
Qy	85	sLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuG1	105
Db	20356	ACTAGAAATCTCGATCCAAATTCAAAGATCTACGCGCTTTGCAATCTCAAAATTTGCT	20297
Qy	105	yTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPhe-----HisG1	123
Db	20296	TCATAA--CAAGAAATTTTAAAGAAAGTGCAAAATATCTTAAATTTGCCAACGACGA	20240
Qy	123	uileTyrPheAsnGlnAlaThrSerGlyValTyrMetCysAlaValAlaIleAlaLe	143
Db	20239	AGTCATGAAGTTTAAAAAATCTGACTCA-----AGTCGAAAT	20201
Qy	143	uGlyTyrLysGluIleTyrLeuSer-----GlyIleAspPheTyrG1	157
Db	20200	TCATATACCGAATTTTATAAATCACTAAAAATCAAAAAACAGAAAAATAATTTTAA	20141
Qy	157	nAsnGlySerTyrAlaPhe-----AspThrLysGlnG1	169
Db	20140	AAATCTCTTAATATGCTTTATTGTTGGCATGCTAGGCTGGAGCACTTTACTGGA	20081
Qy	169	uAsnLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSe	189
Db	20080	AAACATATTGAGTTTAAATCTCT-----	20059
Qy	189	rLysAsnThrAspIleLysAlaLeuGluPheLeuGluLysThr-----	203
Db	20058	-AAAGTAAGTATGGTGAGGTGAGTTTATGAGGGAATCTATCAAGGAAGCTAAAGA	20000
Qy	204	-----TyrLysIleLysLeu-----	208
Db	19999	TTTAGAGATGTTTCATAACTTATACGAAAAAATAATAATCAATTTGATCATCTCT	19940
Qy	209	-----TyrCysLeuCysProAsnSerLeuLeuAlaAs	219
Db	19939	TATTTACACCGACAAAAATTTATTATATATGATTTGTGCC-----ATTATTCTAA	19886
Qy	219	nPheIleGluLeuAlaProAsnLeuAsn-----	228
Db	19885	TTTTTCCCTTAATGCAAAAAATAATAATGCTAGAACCCCTCTTGATATAATTTATC	19826
Qy	229	-----SerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuI1	245
Db	19825	CATATACAGACCAACTTTTAAATCAGTCTTCTCTCTTCTTGCTGATATT-----	19771
Qy	245	eProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLysIleLysI1	265
Db	19770	-----GCTAGTTTATACTTACACTATTTTGAACATTATTAAGGAATATAAAATTAATA	19718
Qy	265	eLysGluAsnValTyrTyrLysLeuIleLysAspLeuLeuArgLeuProSerAspIle	284
Db	19717	TGGTGAAATATTATGATATTACTATGAGGACTTAATGAAAATCCTAATAATGTT	19660
RESULT 118			
AF288092			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

JOURNAL REFERENCE AUTHORS TITLE JOURNAL	Unpublished 2 (bases 1 to 49843) Burger,G. Direct Submission Submitted (18-JUL-2000) Departement de biochimie, Universite de Montreal, 2900, Boul. Edouard-Montpetit, Montreal, Quebec H3T 1J4, Canada
FEATURES	Location/Qualifiers
source	1..49843 /organism="Naegleria gruberi" /organelle="mitochondrion" /mol_type="genomic DNA" /db_xref="taxon:5762" 128..2800 /gene="rnl" 128..2800 /gene="rnl" /product="large subunit ribosomal RNA" 199 /gene="rnl" /note="sub1; polymorphism" /replace="a" 201 /gene="rnl" /note="sub2; polymorphism" /replace="a" 2825..2896 /gene="trnN(guu)" 2825..2896 /gene="trnN(guu)" /product="tRNA-Asn" /note="post-transcriptional RNA editing at 5'end presumably restores base pairing in the acceptor stem" /anticodon=(pos:2857..2859,aa:Asn) 2929..3912 /gene="nad1" 2929..3912 /gene="nad1" /ec_number="1.6.5.3" /codon_start=1 /product="NADH dehydrogenase subunit 1" /protein_id="AAG17776.1" /db_xref="GI:10444210" /translation="MLFFPKNIIIFLLFLVFLLLTIALTLTVERKAMASIORRGEN VGFQFQPIADGKATKEIVLFSNMLFPLAPVTFPFCISILLVWVVPGENKY LIDPFLSLYLLTSSLYVGIIFSGSSNSYAFGLGURSAQMISYEVCIGVFIV MIMTDSNLIIDVINQEVWFIPLFPLMILFFISALAEINRAPFDLPAAELVAG YNVESSITAFALFILAEGNIVLMSLLISIFPWGGLAIPLPFLNLLPFFTPALKVT LHVFIVIRAAFPYRYDQLMKLGWKVLPISVLVYLISFILLNLYIL" 3923..5146 /gene="rps7" 3923..5146 /gene="rps7" /codon_start=1 /product="ribosomal protein S7" /protein_id="AAG17777.1" /db_xref="GI:10444211" /translation="MFNYLLNKKKKKKVVRVNTLFSPLPKKKKKLKLIRINRPNKLNKKY SSDFYFWKYIGDFVISKFINFLVKSQHNRAIKLFFKALLNKKFGNINPILIKY IVLKNVLHKVTKKKVKOKTFYSLRLDPEKQISNTIKIMELIFPKNKKIRLWQS IFLVLNLSILKNNKKTAPFSYDSLNENINYSKVLRTFSLRFFKQVKNKFLRLVYS LKGSVFLMNKVPFIFFYFKKYLGNLNSVSLDPSILHLHSFNKNLYFYLTCKMYRL AILNSIFVDKLQKYREKLKNFFSFRRLKRLHLLNRIDLFRKKYKKRMLLRNKLRLR SKYQHFILTRFFFGWGLAIPLPFLNLLPFFTPALKVT RKIFRVSKYGQIKK" 5243..5313 /gene="trnF(gaa)" 5243..5313 /gene="trnF(gaa)" /product="tRNA-Phe" /note="post-transcriptional RNA editing at 5'end presumably restores base pairing in the acceptor stem; copy number 1"

[illegible]

```
Db 15032 AATAGTTTAAATAAAAAATATATTAGTTTAAATTTA-----TTTAGTAGTGAT 15079
      |||   :::   |||
Qy 130 IleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyr 149
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15080 ATTAATTTTAAATTCATATCTTTCTAGTTTAGGTATT-----TTTGATAATTTAGAA 15130
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 150 LeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLysGlnGlu 169
      ::::   ::::   |||   |||   |||   |||   |||   |||   |||   |||
Db 15131 ATTAATAATTTAAATTTTATTCTAATATTCTCGTTAT---TTTGAT----- 15175
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 170 AsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHisSer 189
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15176 -----TTAGAAATTAAGTAGTATTAGTTAAATAATCATGATTTATTA----- 15217
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 190 LysAsnThrAspIleLysAlaLeuGluPheLeuGlyLysThrTyrLysIleLysLeuTyr 209
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15218 AAACAAACAGATTTA-----TTTTCTTATTACTATAAAACAAATAAGTTGAA----- 15265
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 210 CysLeuCysProAsnSerLeuLeu-----AlaAsnPheIleGluLeuAlaProAsn 226
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15266 -----AATAATAATTTCTTTTTTTATTATCATGGTAACATTATTACCATTAGTTGAAAT 15319
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 227 LeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThrLysAspIleLeuIlePro 246
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15320 GTATCA-----GAACAAATAACAAATTA-----GTTTTTTTACCA 15355
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 247 SerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe----- 260
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15356 AGTTCTAATTTATTTCAGGTTAGAAGATAGTTATATAAATATTTTGTGATTTACAATCA 15415
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 261 -----LysLysIleLysIleLysGluAsnValTyrTyrLys 272
      ::::   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15416 GTAAATCAATCTCAAGATTTTTCAGAAAAAGAAAGTATCAAAATCAGATTTTAT----- 15469
      |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
Qy 273 LeuIleLysAspLeuLeu 278
      :::   |||   |||   |||   |||   |||   |||   |||   |||   |||
Db 15470 ATTTTAAAAACAATTTTA 15487

RESULT 119
LOCUS AB014837 257757 bp DNA linear INV 07-OCT-2002
DEFINITION Plasmodium falciparum 3D7 chromosome 11 section 2 of 8 of the
complete sequence.
ACCESSION AB014837 AEO14186
VERSION AB014837.1 GI:23495936
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 257757)
Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T.,
James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A.,
Kyes S., Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J.,
Angiuoli S., Pertea M., Allen J., Selegut J., Hart D.,
Mather M.W., Vaidya A.B., Martin D.M.A., Fairlamb A.H.,
Fraunholz M.J., Roos D.S., Ralph S.A., McFadden G.I.,
Cummings L.M., Subramanian G.M., Mungall C., Venter J.C.,
Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M. and
Barrell B.
Genome sequence of the human malaria parasite Plasmodium falciparum
Nature 419 (6906), 498-511 (2002)
PUBMED 12368864
REFERENCE 2 (bases 1 to 257757)
Gardner M.J.
Direct Submission
AUTHORS
TITLE Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr., Rockville, MD 20850, USA
FEATURES
source
1..257757
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
```

```
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="11"
repeat_region
257..381
/rpt_type=tandem
305..373
/rpt_type=tandem
763..787
/rpt_type=tandem
1304..1333
/rpt_type=tandem
complement(<1434..>1994)
/locus_tag="PF11_0070"
complement(<1434..1828,1910..>1994))
/locus_tag="PF11_0070"
complement(join(1434..1828,1910..1994))
/locus_tag="PF11_0070"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN35659.1"
/db_xref="GI:23495997"
/translation="MCMVCIIECTEKKKKKKKKKLPIYIHKIDVKNKAAEQNP
WPRIYKGHLFVEKRPQSTLYFYFTILNMRDVTKIHAVLVAYERKSLDFPVLGKK
ANIVYIVNKKPKKIWEPNIKSISFISITAVTNSKMDETIPSTKVTILFISILSTKIV"
1872..1902
/rpt_type=tandem
1926..1960
/rpt_type=tandem
2199..2240
/rpt_type=tandem
2384..2542
/rpt_type=tandem
2582..2650
/rpt_type=tandem
complement(<2647..>4074)
/locus_tag="PF11_0071"
complement(<2647..>4074)
/locus_tag="PF11_0071"
complement(2647..4074)
/locus_tag="PF11_0071"
/codon_start=1
/product="RuvB DNA helicase, putative"
/protein_id="AAN35660.1"
/db_xref="GI:23495998"
/translation="MEINISNITSSNNKERNVHSHIKGLVNTNIVLHENEISLS
DEKYMFPDSCGLVGQFKAREASLFLVDLIRKQKLAGKICILLAGPSGKGALAI
SREINRMPFVFLSGSEVYSENIKKTEVILEAFKRSIHKIKKEKLVEGEVDMVYE
ENECLYSLNKAKQINAIITLKSVKSGKTLRLAPKIHQIAREKIKIGDVIYIETNG
HVKRLGRNDYAKEDYIDFDEYVSLPKGEVHKKEVVOQISLHDIDLANANPTVGSD
ASVLNSVLRPKTEITKRLVEINKTNKLESGLAELPGVLYIDEAHMLDIECFYQ
LNRAIESPLAPILMATNRGICTVKGTONIEPHGIPVDLLDRLLIKTFPYLKEIVQ
ILALRAQTEKINITEBGNYLAKIGIQSSLPAMLLLPSPRLATLEGNTIIDIRHIE
QADLEFMDAKTSAHRVVDQSNKFN"
1459..4258
/rpt_type=tandem
4159..4215
/rpt_type=tandem
4718..4749
/rpt_type=tandem
4922..4978
/rpt_type=tandem
5175..5199
/rpt_type=tandem
5322..5352
/rpt_type=tandem
5350..5397
/rpt_type=tandem
5405..5438
/rpt_type=tandem
5531..5557
/rpt_type=tandem
<5712..>6914
/locus_tag="PF11_0072"
```







```
/product="Chlorodoxin, putative"
/protein id="CAD51844.1"
/db_xref="GI:23505062"
/db_xref="GOA:O81221"
/db_xref="InterPro:IPR006662"
/db_xref="InterPro:IPR006663"
/db_xref="UniProtKB/TrEMBL:O81221"
/translation="MKCYENMLTHINILYSQSIYELKNTGSLNQVSSFTSCNSIIV
KFGAVCKPCNKIKEYFKQLNYYVTVLVDIVDHPKLNQDHNKALPTTEFFYNLN
NEWLVHVEGANQNDIEKAFQKCYCLEAK"
join(10913..11086,11363..114623)
/gene="PFI0795w"
/join(10913..11086,11363..114623)
/gene="PFI0795w"
/feature="Signal anchor predicted for PFI0795w by SignalP 2.0
HMM (Signal peptide probability 0.152, signal anchor
probability 0.803) with cleavage site probability 0.079
between residues 40 and 41
1 probable transmembrane helix predicted for PFI0795w by
TMHMM2.0 at aa 21-43"
/codon_start=1
/product="hypothetical protein"
/protein id="CAD51845.1"
/db_xref="GI:23505063"
/db_xref="InterPro:IPR002885"
/db_xref="UniProtKB/TrEMBL:O81220"
/translation="WGNKMLTKKDKNGKKAIVFLILILIFVDQWSPFPIVYSCFIN
KRSFLVNS IGRNDGPKYIRIDNGKTLLEIGNSKEQNVSKALANDELFLAAPDKK
LSKMSMIGCVNLGCRFNDINMKVIERMKHNIEMOENSLAI CNYTISNNYIV
EYILTINKMIKNTTIRERFVYILVHLDLSFEKNVASORMDYNLDKTKDSLSIG
REIHPHNNIHMIDNNERNLNIKNIDYNNFKYQNMKNLYDITLEQEDQKYLHNLY
LIDIFKKNKIKILYLLKLYLYVYENIYQYIRKKNANEKLITNKEQNISINDT
DORNEKLDKDKFNDGHEDETDNSGEQIDMHKQNTKKNKLLFKLLDILILEQLR
NILELYNYESMINIKYBEETDEERSIYVNVKIKTKLPFKLTENIKNTYNC
KNCENINTYFLSKHIIIVINIIMI THQFNKELIKI KHFSTLNGPVDYTGEDIE
KKETESGGTNOQDVGHNVKEQNEQDNNITNENNEKNSGIMSHPQINGNER
NHTSINEMNKLFPKNTYCLLDGANIGYKNQVNGRPSFLQELIKELIKKKNEX
PLIILPKIYHSEENKERISQOEIICNNNDVLKNTVLTNDNLDNLKKNKSLKDIIE
DSGQHLINSEENKERISQOEIICNNNDVLKNTVLTNDNLDNLKKNKSLKDIIE
STEQNNPEKSKKKNNNMKNNTFYISNIYKCVTHMSVKQPIYIYNTDRMKNLE
YFNFTYILKWKKSISFYKQPVILSEIKEOAGNNAIDVENILQNYLPLFVNLE
NFKLYIDDIIVSKQKQVYHIIKINTPYKTFMSYQKDPFPMGSHNVVVKLCIDFSKI"
complement(join(15222..15364,15455..15582,15715..15804,
15947..16143,16443..16553,16671..16777,16874..16928))
/gene="PFI0800c"
/complement(join(15222..15364,15455..15582,15715..15804,
15947..16143,16443..16553,16671..16777,16874..16928))
/feature="4 probable transmembrane helices predicted for
PFI0800c by TMHMM2.0 at aa 7-29, 88-110, 130-152 and
162-184"
/codon_start=1
/product="hypothetical protein"
/protein id="CAD51846.1"
/db_xref="GI:23505064"
/db_xref="UniProtKB/TrEMBL:O81219"
/translation="MFGCKTHFNILDAFCVGIPIFSFETTYGIFVFPVNDYSIKWHE
PESHACKENVALQIYNVASKGEFTNRMRKQCHSINFLKNLILLITFIVSIIDLIV
VKGVLVLAHPREHETALHKVFCNKLITLICIITNCILLVAPVFFNFSYBEFGK
THYLHAGFYILLADFLTHIIVIFLKRDKTILKVCLELIPWPBKRKLPHVYAEFND
TLISFNEITNYLQGLFQODNINEQLKNFFGTDNYEQIYYVILNKNQKQMMNITPM"
join(18553..20050,20090..26112)
/gene="PFI0805w"
/join(18553..20050,20090..26112)
/gene="PFI0805w"
/feature="No Pfam hit, SMART hit to SM00184, Ring finger"
/codon_start=1
/product="hypothetical protein"
```

```
/protein id="CAD51847.1"
/db_xref="GI:23505065"
/db_xref="GOA:O81218"
/db_xref="InterPro:IPR001841"
/db_xref="UniProtKB/TrEMBL:O81218"
/translation="MTKIQNDIMLPFFSVSSKHCNKNTNVVYVTRTKRRKRKRKKKEK
YCHRCINIKKLYKSKYNNKSNVYIDLKQMRIRKNSYENDYKIFVFKYNNIKSNIIN
SNNIKNYKKYKLESILPFPYNDYSITSMYIVESLIVLTVWSNGIISLFNKHNYHIVSRKYT
NHLIKNLQNCWCSNMIAPNDMDSLVFINLNNIINKGLHKVNIIFLSDNIVKGYNNY
NIDYNNREIDRLSDVFTYTNKLKKKKKKYDNNKDNINVLISDEIMKYNKLNLT
FCINPYNNNNNNNNNNICVLTKRKTSLIYIYEHINSFKDDNINLNIYIESFIF
```

Alignment Scores:	
Pred. No.:	Length: 341050
Score: 124.00	Matches: 74
Percent Similarity: 36.9%	Conservative: 47
Best Local Similarity: 22.6%	Mismatches: 119
Query Match: 7.9%	Indels: 88
DB: 13	Gaps: 20

US-10-734-719-9 (1-291) x PFA929357 (1-341050)

QY	13	LeuLysGluIleAspTyrSerArgLeuProAsn---AspPheAspValPheArg-----	29
DB	13409	ATAAAAGAAATA---TTTAATAGATTAGATTAGTCTAGTATATAAAGAAATCG	13465
QY	30	-----CysAsnGlnPheTyrPheGluAspLysTyrTyr--- 40	
DB	13466	GAAGAAAGAAATGTTTA	13525
QY	41	LeuGlyLysLysCysLysAlaValPheTyrProAsnPhePheGluGlnTyrTyr 60	
DB	13526	TTAGGAAGTTTACGAAAAAGT-----AATAATTATTTTAAATATATATATATAT	13570
QY	61	ThrLeuLysHisLeuIleGlnAsn---GlnGlyTyrGluThr-----GluLeuIleMet 77	
DB	13571	TATATAGATTATCTAAACATTTTTCATAATATCAACATTTGAATCATAATATATTATT	13630
QY	78	CysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheVal----- 92	
DB	13631	ATAGATAATATATAAT-----ATTATGGATAATAATCTAATTTACAATAACACGGAA	13681
QY	93	-----LysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyr 106	
DB	13682	GPATTATATGTCATAATAATGTGATTATATGATAAGATAGTAGTATATTATGGTTTATAGTA	13741
QY	107	AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPhe 126	
DB	13742	AGTGAAGAAATATGNAATCGAGAA---AATAATATATATCTATGATGAACATAATATAT	13798
QY	127	AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleLeuGlyTyr--- 145	
DB	13799	AATGATAAAATTAAGAAAAAT-----AATGCTAATGTAAATTTATTGACAAATATAAA	13852
QY	146	-----LysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsn 158	
DB	13853	AATAATATGAAATTTTATAGATAAGAAATTTTAAATAATAATAATAATATTTTATGAT	13912
QY	159	GlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPhe 178	
DB	13913	ACATATATGATGATGATGATGGACAAACATATATTTGGTTAAT----- 13957	
QY	179	LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGlu 198	
DB	13958	-----TCTCATGAAGAAAAATAA----- 13975	
QY	199	PheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAla 218	
DB	13976	-----GAAAGCAAGTCAACAAGAAATATTATTGT---AACAAATATGATGATTATTAAG	14026
QY	219	AsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsn 238	
DB	14027	ATAAAGTAATAATACATTAGATAAT---AATGATAATTTTATTAATTAATAAATAAGAAATCG	14083

QY 239 TyrThrLysAspIleLeuProSerSerGluAlaTyrGly-----Lys 253  
Db 14084 ---TTAAGGATATTATGATGAAATCCACAGACAGAAAATAACCTGAAAAGGAAAA 14140  
QY 254 PheSerLysAsnIleAanPheLysLysIleLys-----IleLysGluAsnVal 269  
Db 14141 AGTAAAAAATAAATAATAATAATATGAAGATAATACATTTTATTAGTAACATA 14200  
QY 270 TyrTyrLysLeuIleLysAspLeuArgLeuProSerAspIle----- 284  
Db 14201 TACATAAAATGCTGTACACATATGAAAAGTGTTCAAAAACCTATTATTATTACTAAT 14260  
QY 285 -----LysHisTyrPhe 288  
Db 14261 GATAAAATGAAAAATCATTATTTT 14284  
RESULT 121  
AF494278/c 131183 bp DNA circular PLN 05-DEC-2005  
LOCUS  
DEFINITION Chaetosphaeridium globosum chloroplast, complete genome.  
ACCESSION AF494278 AF393581  
VERSION AF494278.1 GI:22416897  
KEYWORDS  
SOURCE  
ORGANISM  
chloroplast Chaetosphaeridium globosum  
Chaetosphaeridium globosum  
Eukaryota; Viridiplantae; Streptophyta; Coleochaetophyceae;  
Coleochaetales; Chaetosphaeridiaceae; Chaetosphaeridium.  
REFERENCE  
AUTHORS Turmel,M., Otis,C. and Lemieux,C.  
TITLE The chloroplast and mitochondrial genome sequences of the  
charophyte Chaetosphaeridium globosum: Insights into the timing of  
the events that restructured organelle DNAs within the green algal  
lineage that led to land plants  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (17), 11275-11280 (2002)  
PUBMED 12161560  
REFERENCE  
AUTHORS 2 (bases 1 to 131183)  
TITLE Turmel,M., Ehara,M., Otis,C. and Lemieux,C.  
Direct Submission  
Submitted (20-JUN-2001) Biochimie et Microbiologie, Universite  
Laval, Pavillon Charles-Eugene Marchand, Quebec, Quebec G1K 7P4,  
Canada  
REFERENCE  
AUTHORS 3 (bases 1 to 131183)  
TITLE Lemieux,C., Otis,C. and Turmel,M.  
Direct Submission  
Submitted (22-MAR-2002) Biochimie et Microbiologie, Universite  
Laval, Pavillon Charles-Eugene Marchand, Quebec, Quebec G1K 7P4,  
Canada  
COMMENT On Aug 22, 2002 this sequence version replaced gi:21635202.  
FEATURES  
Location/Qualifiers  
1. .131183  
/organism="Chaetosphaeridium globosum"  
/organelle="plastid:chloroplast"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:96477"  
/cell\_line="M1311 (collection of M. Melkonian, University  
of Cologne)"  
join(37526..38253,127..1307)  
/gene="rps12"  
/trans splicing  
join(37526..37639,1050..1307)  
/gene="rps12"  
/trans splicing  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein S12"  
/protein\_id="AA096595.1"  
/db\_xref="GI:22416995"  
/translation="MPTIQILRNREPFIENRTKSPALRSCPQRRGVCTRYVTPPK  
PNSALRKARVRLTSGFEVTAYIPGIHNLQEHSVLVGRGRVDLPQVRYHVRGTL  
DSVGVKDRHGRSKYGAKPK"  
complement(join(126058..126095,125360..125394))  
/product="trna-Ala"

trna

misc\_feature

intron

gene

CDS

exon

gene

CDS

stem\_loop

gene

CDS

exon

intron

exon

gene

/note="codons recognized: GCN"  
/anticodons(pos:126060..126062,aa:Ala)  
complement(join(126748..126784,126143..126177))  
/product="trna-Ile"  
/note="codons recognized: AUU"  
/anticodons(pos:126750..126752,aa:Ile)  
1..88682  
/note="Region: large single-copy region (LSC)"  
127..1049  
/gene="rps12"  
/note="3' part; domains IV, V and VI of a trans-spliced  
group II intron"  
/number=1  
196..939  
/gene="orf247"  
196..939  
/gene="orf247"  
/codon\_start=1  
/transl\_table=11  
/product="putative intron-encoded protein"  
/protein\_id="AA096587.1"  
/db\_xref="GI:22416987"  
/translation="MLRAISNDKLTWKINKWKIENNLFKIOTRIYKAKLKQKIKLK  
QNKELMSKNFELRKIQILFNSKDYIKNOFFEEITKEFLIKIHSPSWTKILE  
NNINQINILRNKDDYNNILNKKDNFFIKILFSFFIKQNFINSYIWFINTK  
FEKDLAVLNTQOFPSPYKEKNKPKQIKNVLNKSINLIKDIYIKENQNKFKQVI  
LITKTDYKKFINSKMTAKINKELTNF"  
1050..1307  
/gene="rps12"  
/number=2  
1336..1803  
/gene="rps7"  
1336..1803  
/gene="rps7"  
/codon\_start=1  
/transl\_table=11  
/product="ribosomal protein S7"  
/protein\_id="AA096578.1"  
/db\_xref="GI:22416978"  
/translation="MSRRKNAEKRIVNDPVPVRNLVNLVNRILKNGKSLAYRILY  
KMTAKITKTQKNFLFILQAIRKATPKVAKRVGGSTYQVPEIQSSQKALAIR  
WLLTSAKKRAGRNFIKLSNEIIDTARDTGNTIRKKBETHRAESNRAFAFR"  
1801..1885  
1953..4034  
/gene="ndhB"  
join(1953..2678,3258..4034)  
/gene="ndhB"  
/codon\_start=1  
/transl\_table=11  
/product="subunit 2 of NADH-plastoquinone oxidoreductase"  
/protein\_id="AA096515.1"  
/db\_xref="GI:22416915"  
/translation="MEIQNNLYGIDLYTILPESILIFCLTTLIIDLSLDSKNSWII  
YLTNTIGILSGFLCLOWNGIIPVTPPFSFKGDSFIARFCIIIASLLSLSDIYI  
KRAGVOLMEFVIFLLGATIGMFLCGANDLITIFTSLECLSSVLLAGYSKQDIRSN  
EAKNLYLVGGASNAIAYGFSMLYGLSGKILSEIVDGLIFADVPNPLIKWITLTC  
EQFLQLIACLTMIVGNLVAITQTSMKRMLAYSSISQAGYLMIGIISSTNDGIASSL  
VYMLIYIPMLNAGFCVILFGRTGDIQIRDFSGLYLKDPLMASCLTIFLLSLGGIPP  
PAGFGKIYLFWSGWQAGLYLITFVGLLTSVISIYYLRIIKIMFVREAKEFSSVYKN  
YVIPNSILLPOSSVETAMIVCMIASSVNGIAINPIQVAQKLTILTIPTFI"  
1953..2678  
/gene="ndhB"  
/number=1  
2679..3257  
/gene="ndhB"  
/note="group II intron"  
/number=1  
3258..4034  
/gene="ndhB"  
/number=2  
4119..4199  
/gene="trnL(caa)"

```

4119. .4199
/gene="trnL(caa)"
/product="trnA-Leu"
/notes="codon recognized: UUG"
/anticodon=(pos:4152..4154,aa:Leu)
4309. .4413
/gene="psbM"
/gene="psbM"
/codon_start=1
/transl_table=1
/product="M protein of photosystem II"
/protein_id="AAM96549.1"
/db_xref="GI:22416949"
/translation="MEVNILAVIATLFLVLIPTAFLLILYVKTESAGS"
4534. .5486
/gene="ycf66"
/join(4534..4639,5131..5486)
/gene="ycf66"
/inferences="non-experimental evidence, no additional
details recorded"
/codon_start=1
/transl_table=1
/product="hypothetical chloroplast RF66"
/protein_id="AAM96586.1"
/db_xref="GI:22416986"
/translation="MIRMEIGPSTFVGICLIIMGIILYLLTKQPEVSKVDLLFSSI
GLUGGLIFQGWRLDPIILLCQILCSGTAFVLESLYLRNQSUKYISKQVQDDKKV
TLNNTIKLTSEKSDINIKYFKIKALIEAKVPPEYTHNISYDLLKKKA"
4534. .4639
/gene="ycf66"
/number=1
4640. .5130
/gene="ycf66"
/notes="group II intron"
/number=1
5131. .5486
/gene="ycf66"
/number=2
complement(5570..5659)
/gene="petN"
complement(5570..5659)
/gene="petN"
/codon_start=1
/transl_table=1
/product="cytochrome b6/f complex subunit VIII"
/protein_id="AAM96531.1"
/db_xref="GI:22416931"
/translation="MDLITITWASVMVAFTFSLVWGRSGL"
complement(5827..5898)
/gene="trnC(gca)"
complement(5827..5898)
/gene="trnC(gca)"
/product="trnA-Cys"
/notes="codons recognized: UCG"
/anticodon=(pos:5864..5866,aa:Cys)
6085. .9297
/gene="rpoB"
6085. .9297
/gene="rpoB"
/codon_start=1
/transl_table=1
/product="beta subunit of RNA polymerase"
/protein_id="AAM96566.1"
/db_xref="GI:22416966"
/translation="WTIQNTQTEKFTLPDLGKIQLQNFYNFLTKIYKELKAFPKIYD
SEHNFEPKYPDEFVLPILNERESIYKSTIYELVTAELNCKIKQTKQKQKVL
GHPLITTSNGFIINGSRVQSILRSPGFSGSDNSKCIYATYIISESGRLK
LEIDDRKLIWARIKKRVSAILLLSMGILKSDITRSIEHPAILDNLKKNVLSNQ
KATLEYIKOLFSAAGDPNFDNLSLQRTFFQORCEIGIGRNLKGLKLNIPDE
EFLLPEDILATNYLKLKSLGIGSLDDIDHLMKRVKSIISDFLSEOLKNALKIENT
IEQIKNISKRKLILTPKLSGILLAFFKDFGSHPLSLOFLQTNLSLTLVHKRI
SSUGPGGLTRTASFHVRDIIHPSHYGRICPIETSEGMNAGLVASLASNATINTMGFIE

```

Alignment Scores:

Pred. No.:	0.0012	Length:	131183
Score:	123.50	Matches:	80
Percent Similarity:	29.9%	Conservative:	55
Best Local Similarity:	17.7%	Mismatches:	138
Query Match:	7.9%	Indels:	178
DB:	4	Gaps:	14

US-10-734-719-9 (1-291) x AF494278 (1-131183)

QY	2	LysLysValIleIleAlaGlyAsnGlyProSerLeuLysGluIleAsp	-----	17
QY	79919	AAATCAACAATTTTAAAGGAGAATCAAAACAGTTTTCAAATCAATAACAATATCTAA		79860
QY	18	-----TyrSerArgLeuProAsnAppPheAspValPheArg		29
QY	79859	AAAAAGCAATTTTCTCTTAAAGTATTCAGAAAAGTTTAAATCTTTCAAGATATTTAGA		79800
QY	30	CysAsnGln-Phe-----TyrPheGluAspLysTyrTyrLeuGlyLysCysLysAl		47
QY	79799	AAAAAGCGTTTTCAGAAAATTTTTCGTAATAAAAAATTTTATGATTTTATACAAAGTT		79740
QY	47	aValPheTyrThrProAsnPhePheGluGlnTyr-----		60
QY	79739	TCTATTTTATTAATAAATATATTTATCAACAAATTTTAAACGAATATTTTATAAATA		79680
QY	60	rThrLeuLysHisLeuIleGlnAsnGln-----GluTyrGluThrGluLe		75
QY	79679	TCGATTTAAATTTTATATACAAAAACAAACATAGATTTTGAATCAAAAAACAGAGT		79620
QY	75	uileMetCysSerAsnTyrAsn-----		82
QY	79619	ATTAATAGTAATAATTAATTTTATAAAAAATTTTAAAGAAATTTGAAAAAGCAACA		79560
QY	82	-----		82
QY	79559	ACTGAGTCAATTCAAATTCATCTCTTTTACAAATTTATATATATTGAAGAAATAG		79500
QY	83	-----GlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTy		98
QY	79499	AAATTTCTTATAATATCCATCTCGAAAAATAATAGAAATTTTAAAAACTATTACTACT		79440
QY	98	rPheProAspAlaHisLeuGlyTyr-----		106
QY	79439	TACAGACAATAATTAATTTTATTTAAAGATAATTTCTCTTACAAAAACGATAATTT		79380
QY	106	-----		106
QY	79379	TTTAATTTCAAAAAAAATTTAACGGATTCAAAAAATTTATATAAAAAATTAATAAGTT		79320
QY	106	-----		106
QY	79319	ACAAAAAGATAATTTTCTGAGGTTTGTAGAGAACAAAAACACCATTTATTTTAAACAA		79260
QY	107	-----AspPhePheLysGlnLeuLysGlu-----		114
QY	79259	AAATTTAATAAATACTGACCTCAACTAATTTTATTTTACAAAAAGAGAAGAAATC		79200
QY	115	-----PheAsnAlaTyrPhe-----Ly		120
QY	79199	AAAAAAAATTTTCAAGATTTCATTTGAATTTAATTAATTTTCTCTATATTTTAGGAA		79140
QY	120	sPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAlaValAl		140
QY	79139	ATTTAAAGCGTAAATTTAATAGTTTAAATCAACCATTAATTCATATAATAAAAAAT		79080
QY	140	allelaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySe		160
QY	79079	ATTAACCTTTTAATATAAAAAATGATAAGATAAAAAATGTAGTTTATTTAAAAAT--		79023
QY	160	rSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuAlaProAspPheLysAs		180



DEFINITION Sequence 107 from Patent WO0168911.  
ACCESSION AX251846  
VERSION AX251846.1 GI:15985201  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with the cell cycle  
JOURNAL Patent: WO 0168911-A 107 20-SEP-2001;  
EpiGenomics AG (DE)  
FEATURES  
Location/Qualifiers  
1. 6129  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
ORIGIN  
Alignment Scores:  
Pred. No.: 7.22e-05 Length: 6129  
Score: 123.00 Matches: 92  
Percent Similarity: 33.6% Conservative: 39  
Best Local Similarity: 23.6% Mismatches: 108  
Query Match: 7.9% Indels: 154  
DB: 2 Gaps: 18  
US-10-734-719-9 (1-291) x AX251846 (1-6129)  
QY 31 AsnGlnPheTyrPheGluAspLysTyr-TyrLeuGlyLysLysCysLysAlaValPheTy 50  
DB 2192 AATCGCTTCCTTTTCAACAAATAAATACGATTAACTAAAAA---AAATCCTACACACA 2136  
QY 50 rThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuLysGlnAsnGlnG 70  
DB 2135 CACAAAACATCTATATATATAAAATAATAC-----AACTTAATTC-AACTCAACA 2086  
QY 70 uTyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAs 90  
DB 2085 A---AAATTCATATAAACAATATAAATAACATCAACAATACCTCACTCGAAATATCAA 2029  
QY 90 nPhe-----ValLysThrPheTyrAs 97  
DB 2028 CATTCCTACTATTATAATAATTTTATAATTAATAAATAACATCACCTCAAAACCATTTATAT 1969  
QY 97 pTyr-----PhePro-----AspAlaHisLe 104  
DB 1968 CTACTATTATATACATACCAACAAATAAATAATCATACTAACTACATACAAACAAAT 1909  
QY 99 -----PhePro-----AspAlaHisLe 104  
DB 1908 AAAATTTCCAAACATACCAACCAACCAACCAATCAATCAATATCGTATCTCTTAATAT 1849  
QY 104 uGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluL 124  
DB 1848 TTAACACTTCATTTCACATTAATAAATAAATAAATAATTAATTTTCTTTTACCACTTT 1789  
QY 124 eTyrPheAsnGlnArgGlyThrSerGlyValTyrMetCysAlaValAlaLeuAlaLeuG 144  
DB 1788 AAATCATATCAATATTTTCGAAAAATCTATAC----- 1755  
QY 144 yTyrLysGluLeuTyrLeuSer-----GlyIleAspPheTyrGlnAsnGlySerSerTy 162  
DB 1754 -TTTAAACAATATATTAACCCCTCATCAATAAATTTTATTTCAACAAAAAATA 1696  
QY 162 r-----AlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAl 175  
DB 1695 CTTCATATATTCTAACTTAATTTTAAAAAACCGAACGCAATAACTC-ACGCCTATAAT 1637  
QY 175 aproAspPhe---LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAsp 194  
DB 1636 CCCATCACTTTTAAAAAACCGAA-----ATAAACGAATCAGGAATAAAAAAAT 1589

QY 194 eLysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCys----- 210  
DB 1588 TAAACCATC-----CTCACTAACACGATAAAACCACTCTCTACTA-AAAATATATAAA 1536  
QY 210 ----- 210  
DB 1535 AATTATCCGAACGATATATCGAACGCTTATATATCCCAACTACTCGAAAACTAAAAACA 1476  
QY 211 -----LeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaPro----- 225  
DB 1475 AATAACGTAACCAACCAATAAAGAACTTACATAAAGCAATTAACCACTACGACCTACATC 1416  
QY 226 -AsnLeuAsnSerAsnPheIleIleGln-GluLysAsnAsnTyrThrLysAspIleLeu 245  
DB 1415 CAACCTTAACGACAAATAAATAAATACTCAATCTCAAAAAAATAAATAAATAAATAA 1356  
QY 245 leProSerSerGluAlaTyrGlyLysPheSerLysAsnIle----- 258  
DB 1355 TAAATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1296  
QY 259 -----AsnPhelyLysIleL 264  
DB 1295 CAAACCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1236  
QY 264 ysile-----LysGluAsn----- 268  
DB 1235 AATTCGCCACCTCCAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1176  
QY 269 --ValTyrTyrLys-----LeuIleLysAspLeuLeuArgLeuP 281  
DB 1175 AAACATATTTAACATTTTAAACCTTATAATTTCTTATATATTCCTTATATAATATTC 1116  
QY 281 roSerAspIleLysHisTyrPheLys 289  
DB 1115 CATCAATTCCTAAACATATTTTCAAAA 1090  
RESULT 124  
AX344232/c 6129 bp DNA linear PAT 01-FEB-2002  
LOCUS AX344232  
DEFINITION Sequence 79 from Patent WO0200926.  
ACCESSION AX344232  
VERSION AX344232.1 GI:18492120  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM other sequences; artificial sequences.  
REFERENCE 1  
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.  
TITLE Diagnosis of diseases associated with signal transduction  
JOURNAL Patent: WO 0200926-A 79 03-JAN-2002;  
EpiGenomics AG (DE)  
FEATURES  
Location/Qualifiers  
1. 6129  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
ORIGIN  
Alignment Scores:  
Pred. No.: 7.22e-05 Length: 6129  
Score: 123.00 Matches: 92  
Percent Similarity: 33.6% Conservative: 39  
Best Local Similarity: 23.6% Mismatches: 108  
Query Match: 7.9% Indels: 154  
DB: 2 Gaps: 18  
US-10-734-719-9 (1-291) x AX344232 (1-6129)  
QY 31 AsnGlnPheTyrPheGluAspLysTyr-TyrLeuGlyLysLysCysLysAlaValPheTy 50  
DB 2192 AATCGCTTCCTTTTCAACAAATAAATACGATTAACTAAAAA---AAATCCTACACACA 2136



```
QY 162 r-----AlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAl 175
Db 1695 CTTCACTATATTTCTAACTTAAATTTTAAACCGCAACGCAATAACTC-ACGCCATATAAT 1637
QY 175 aProAspPhe-----LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIl 194
Db 1636 CCATCACTTTTAAACCGAA-----ATAACGAATCAGCAATAAAAAAAT 1589
QY 194 elysAlaLeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCys----- 210
Db 1588 TAAACCATC-----CTCACTAACACGATAAAACCAATCTCTACTA-AAAATATAAAA 1536
QY 210 ----- 210
Db 1535 AATTATCGAAGCTATATCGAAGCGCTATATCCCACTACTCGAAAAAATAAACAAAA 1476
QY 211 -----LeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaPro----- 225
Db 1475 AAATAACGTAAACCCAAATAACGAAACTTTACAATAAAACCGAATTCACGCCACTACACTC 1416
QY 226 -AsnLeuAsnSerAsnPheIleGln-GluLysAsnAsnTyrThrLysAspIleLeuI 245
Db 1415 CAACCTAAACGCAAAATAAACTCAATCTCAAAAAAATAAAAAAATAAAATTTCTAA 1356
QY 245 leProSerSerGluAlaTyrGlyLysPheSerLysAsnIle----- 258
Db 1355 TAAATCAATAATAATAAAATTTAAATAATTTCAAAAAAATACCAATAATTATAAAAAATTAA 1296
QY 259 -----AsnPheLysLysIleL 264
Db 1295 CAAACTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1236
QY 264 ysIle-----LysGluAsn----- 268
Db 1235 AATCCCCACCTCCCAACAAAAATAACAATAATATATATTCCTACCCCTTTAAATTCCT 1176
QY 269 --ValTyrLys-----LeuIleLysAspLeuLeuArgLeuP 281
Db 1175 AAACATATTTTAAACTTTTAAACTTATAATATTTCTTATATATATTCCTTATATAATATTC 1116
QY 281 toSerAspIleLysHisTyrPheLys 289
Db 1115 CATCATTTCTAAACATATTTCAA 1090
RESULT 126
AE014075_40
WPCOMMENT
Sequence split into 53 fragments LOCUS AE014075 Accession AE014075
Fragment Name Begin End
AE014075_00 1 110000
AE014075_01 100001 210000
AE014075_02 200001 310000
AE014075_03 300001 410000
AE014075_04 400001 510000
AE014075_05 500001 610000
AE014075_06 600001 710000
AE014075_07 700001 810000
AE014075_08 800001 910000
AE014075_09 900001 1010000
AE014075_10 1000001 1110000
AE014075_11 1100001 1210000
AE014075_12 1200001 1310000
AE014075_13 1300001 1410000
AE014075_14 1400001 1510000
AE014075_15 1500001 1610000
AE014075_16 1600001 1710000
AE014075_17 1700001 1810000
AE014075_18 1800001 1910000
AE014075_19 1900001 2010000
AE014075_20 2000001 2110000
AE014075_21 2100001 2210000
AE014075_22 2200001 2310000
AE014075_23 2300001 2410000
```

```
AE014075_24 2400001 2510000
AE014075_25 2500001 2610000
AE014075_26 2600001 2710000
AE014075_27 2700001 2810000
AE014075_28 2800001 2910000
AE014075_29 2900001 3010000
AE014075_30 3000001 3110000
AE014075_31 3100001 3210000
AE014075_32 3200001 3310000
AE014075_33 3300001 3410000
AE014075_34 3400001 3510000
AE014075_35 3500001 3610000
AE014075_36 3600001 3710000
AE014075_37 3700001 3810000
AE014075_38 3800001 3910000
AE014075_39 3900001 4010000
AE014075_40 4000001 4110000
AE014075_41 4100001 4210000
AE014075_42 4200001 4310000
AE014075_43 4300001 4410000
AE014075_44 4400001 4510000
AE014075_45 4500001 4610000
AE014075_46 4600001 4710000
AE014075_47 4700001 4810000
AE014075_48 4800001 4910000
AE014075_49 4900001 5010000
AE014075_50 5000001 5110000
AE014075_51 5100001 5210000
AE014075_52 5200001 5231428
Continuation (41 of 53) of AE014075 from base 4000001 (AE014075 Escherichia coli CFT073)
```

```
Alignment Scores:
Pred. No.: 0.00117 Length: 110000
Score: 123.00 Matches: 71
Percent Similarity: 36.7% Conservative: 41
Best Local Similarity: 23.3% Mismatches: 81
Query Match: 7.9% Indels: 112
DB: 15 Gaps: 17
```

US-10-734-719-9 (1-291) x AE014075\_40 (1-110000)

```
QY 24 AspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeu----- 41
Db 21399 GAATTTGACTTCTATCGAATAAATAGATTGGTAGATGGGAAAGTATTATATATACAA 21458
QY 42 -----GlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGlu----- 57
Db 21459 CTAAGAGGCGCTTTTATCGGCTATTTCATCACCAG-----ATCTTTGAAAAACCAA 21509
QY 58 -----GlnTyrTyrThrLeuLysHisLeuIleGln 67
Db 21510 CTGGATACAGAAACAATAACTGTTAGTTATGAATTTATCTCTGCCAACATCTTATCACT 21569
QY 68 AsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGlu 87
Db 21570 AACACCGAGTCT---AGCTATTGGGCACCTCCCTGAAAAATTATACCCGCTGTTTATCCA 21626
QY 88 Asn----- 88
Db 21627 AATTCAAAAAACAAACAATAATGTTTAAACCGTTAAACAGCAAGCTGTTGGCAGG 21686
QY 89 -----GluAsnPheValLysThrPheTyr 96
Db 21687 CTACTTGCTGCTGGTGGCGGTATACAATGGGAACATTCGAAGGATTTCAGAGATACCTCGCGAA 21746
QY 97 AspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsn 116
Db 21747 AAACCTGGGTGGAGATGCAATCAAAAGGCTATGAT-----CAATA----- 21785
QY 117 AlaTyrPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMet 136
Db 21786 -----CTAAATGAAAAA---ACACGGCGCATAGCGGATA 21815
```



QY	137	CysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr	156	
Db	21816	GCAACAGCATCTATCTTTTAAACAACGCT	21863	rRNA
QY	157	GlnAsnGlySerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAlaPro	176	rRNA
Db	21864	ACAGAAATAAATAGTTAC	21896	rRNA
QY	177	AspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAla	196	rRNA
Db	21897	-----GGCAACAACAACTCTTGAT-----AGT	21920	gene
QY	197	LeuGluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeu	216	CDS
Db	21921	ATAGACATAATAGAA-----ATAATATACATTAAGAGACCTTCAAAAGAC	21965	
QY	217	LeuAlaAsnPhe---IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlu	235	
Db	21966	TTAGCTAAGCTTACGAAGAGTTAATAAAATGTAAGAAAAATTTTCTTATCAAA---	22022	
QY	236	LysAsnAsnTyrThrLysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSer	255	
Db	22023	-----CTTGCAAAAACCTCCGAAGCATCTGGAAGATTCAAC	22058	gene
QY	256	LysAsn-----IleAsnPheLysLys-----IleLysIleLysGluAsnValTyrTyr	271	CDS
Db	22059	GCCGAAGACCTTTTAAAGATGAGAAAGGCGAATGCTTCTTAAATTATATGATACCAT	22118	
QY	272	LysLeuIleLysAsp	276	gene
Db	22119	AAACTATCTCTAGAT	22133	CDS
RESULT	127			
LOCUS	AF083031		174133 bp	DNA linear
DEFINITION	Guillardia theta nucleomorph chromosome 3, complete sequence.			
ACCESSION	AF083031			
VERSION	AF083031.2		GI:13794268	
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				







```

/db_xref="UniProtKB/TrEMBL:Q7YZ72"
/translation="MSQNQSKDETTKTRHYHGLISLQNGRGVIKSQSIAPQNSKWSN
INSKKEIGDSNATPGMTTVYVGKIPAQVNTTHIRKLISECTIILRWSQEDPTTKK
LSFGCEFDSPGCVINAVSALNGIKISGGQLLVKQHQHDKIEIAKWQSNRIINELLKH
RGDCITETIINELSDSKLRISIQRLLAGMTFELDSQEQTSTEGHKNKPEKKQLDLD
ERNKRISYNSLDSGSISSOMHKSGLVIGFPONPKELRMKRFEEKNKPEKKQLDLD
DELYELVKLPKWDISNDITFRYLDLLOESVHNPKFISSSNKKTFDMISELIQSDK
ELIQOLDYSHSDYNDISYIHEDSANSKRAESISFSFNPCHKYFENILSPNKSII
KKAANTLTGNTQNSVFNNTNMEKISINLPKQHQFENPISLENNINNNNS
SSSSILNLMKDINHILNSKEEVSAENWKNLKSQFENYKEYVELEMKGIIIGESDPN
ALMTITNFIIEKSSKNLIQETINILIQILDTAESPLRGLFKILDG"
complement(1607..1822)
/locus_tag="IMB.630"
/note="pfam match to entry PF00076 rrm, RNA recognition
motif. (a.k.a. RRM, RBD, or RNP domain), score 43.1,
E-value 4.1e-10"
2363..2370
/note="(a)8"
2480..2489
/note="(t)10"
2580..2587
/note="tgcatgca"
2623..2634
/note="(tttg)3"
2681..2670
/note="(a)10"
2842..2850
/note="(a)9"
2897..4387
/locus_tag="IMB.631"
2897..4387
/locus_tag="IMB.631"
/note="IMB.631, predicted protein, len = 497 aa, unknown;
predicted pi = 4.7528; contains no predicted TM helices;
some similarity to Q9D567, 493051la2irik protein (425 aa,
Mus musculus, EMBL: AK015746, BAB29954); Fasta scores:
E() : 0.00049, 23.438% identity (27.439% ungapped) in 384 aa
overlap, (aa 86-464 of IMB.631, aa 89-421 of Q9D567)"
/codon_start=1
/product="hypothetical predicted protein, unknown
function"
/db_xref="CAD98255.1"
/db_xref="GI:32399015"
/db_xref="GOA:Q7YZ71"
/db_xref="UniProtKB/TrEMBL:Q7YZ71"
/translation="MNHVTAFLERMDLRSKNYIERNGKGLFLLRKTATQISLKNQ
LFVTEYTKGTLTFEELNRSFPICLLPERRVRSKSDGIVVRKFRRCRDHCPEI
FSQPLTSMYAKMFDCYGFISPKFINTCKILYTFQISSELYINPMERVITYDFI
LYWTKNSIRENSETKGLVROFSINHFNENFYANYIERLIIFSGSKNSIIRQDLI
SSQDFARLELRADTSESNPNPQLIFQIFEFYRKSINTEQNGQFNEDSDQ
NQNLVGNFEDKFLPMLQQLMDFYIIPFIVYRIFEVHSENCSITDSNGQRT
AIGIKTVKVLFSLEFRESPOSINWFWNILDNETGYIDQSVISCFWKSVMVASIEQIY
SIESLPLFEHISDSIFDMICPELNGKITKKAFKDSPMASTVISYLCDPKAFIAIEMR
BEIIAQNASRNLENDQISLDLEKPSYNDSESNSTIDCSQESISVT"
3501..3509
/note="(t)9"
3959..3970
/note="(caat)3"
4389..4403
/note="(t)15"
4491..4500
/note="(a)10"
4535..4544
/note="(a)10"
complement(4591..6324)
/locus_tag="IMB.633"
complement(4591..6324)
/locus_tag="IMB.633"
/note="IMB.633, predicted protein, len = 578 aa, unknown;
predicted pi = 6.5934; contains Pfam match to entry
PF00043 GST_C, Glutathione S-transferase, C-terminal
domain; contains no predicted TM helices"
/codon_start=1
/product="hypothetical predicted glutathione S-transferase
domain protein, unknown function"
/protein_id="CAD98256.1"
/db_xref="GI:32399016"
/db_xref="InterPro:IPR004046"
/db_xref="UniProtKB/TrEMBL:Q7YZ70"
/translation="MNNITNSWTSGLANLNMGRNRQENMTGEANIIPNOGNDPHYWGV
QGNGRVSQNLHGVSPKDIYQRGNFVNPMDYQSFSPISPNFTTAAYSISGMKQI
VSTIDSGVAVHYDLGKYVGLGGAIRFEMCKQIKHSFVNISLNDENINIGYSSSV
DQSGSPPLIIHNKKHFLFTVTLRIAKKIEGYGDAYRDYALDIFSETLSWRN
ELLALIEKISESKQFQTKKEATETDLSNRMTNFSRNKIKIISDTKTSAAVSDD
NYTKYLKRRKFKFSVEALISNFSANFPFLKGEAIALDSTAESKEVPRQLCNCPY
YSELFLFSIIYDDSIILKENSKFLNENDVSTKLEEPKHLHDLFQAVMSPYLITQWY
KWEQSDATRESNEHTSNKATNTTNNRVQTESNGVRAATPKKTQDIESTIYNTSIP
INNEPHTNTITYELAPKSOVIIPNNSNNQORVEIPNIYHOTCSSNPVSTNTFNAI
RISNYETNFRPSNPQPSVVNRIANFSQVNNLPFKMGQAQQIISTARQFNNSYSDPQNP
QYENFAPSTFR"
4855..4862
/note="(at)4"
4989..4996
/note="(ta)4"
complement(5167..5226)
/locus_tag="IMB.633"
/note="pfam match to entry PF00043 GST_C, Glutathione
S-transferase, C-terminal domain, score 10.8, E-value
0.0062"
6630..6637
/note="(t)8"
6648..6655
/note="(t)8"
6806..6813
/note="tgcatgca"
6850..6857
/note="(a)8"
7004..7012
/note="(a)9"
7089..7099
/note="(a)11"
7176..9242
/locus_tag="IMB.635"
7176..9242
/locus_tag="IMB.635"
/locus_tag="IMB.635"
/note="IMB.635, predicted protein, len = 689 aa, probably
aminoacyl-tRNA synthetase; predicted pi = 6.6193; contains
Pfam match to entry PF00587 tRNA-synt_2b, tRNA synthetase
class II core domain (G, H, P, S and T); contains Pfam
match to entry PF03129 HGRP anticodon, Anticodon binding
domain; contains no predicted TM helices; good similarity
to SYEP HUMAN, bifunctional aminoacyl-tRNA synthetase
includes: glutamyl-tRNA synthetase (EC 6.1.1.15) (1440 aa,
Homo sapiens, EMBL: X07466, CAA30354); Fasta scores:
E() : 1e-107, 55.233% identity (56.548% ungapped) in 516 aa
overlap, (aa 180-688 of IMB.635, aa 930-1440 of
SYEP HUMAN)"
/codon_start=1
/product="aminoacyl-tRNA synthetase, probable"
/protein_id="CAD98257.1"
/db_xref="GI:32399017"
/db_xref="GOA:Q7YZ69"
/db_xref="InterPro:IPR002314"
/db_xref="InterPro:IPR004154"
/db_xref="InterPro:IPR004499"
/db_xref="InterPro:IPR006195"
/db_xref="InterPro:IPR007214"
/db_xref="UniProtKB/TrEMBL:Q7YZ69"
/translation="MSVDEKLYSEVLDQSLKSLINYSKFDHATPNMESKVEVLKQA
EKHNTDFAKNLLIKSGNEGELFFVLAHETDTKMKNLGOIFGVSGNKLRLADSDVLND
TLKVRKGLTPLUSLIFPEKSGGQIVTDECLDKKVFVPLTNTSESIIHNDIVKFAE
SCGWKWFNSMDLNEEQPAGKPEESKESLIGITADKITSFADWYSQIVKSEMIIE
YDTSGCCVILRPWSYFIWETIQSVFDQKHQVQNAFYPIFVTKKLETKHDSVGF
SPEVAWTKSGKSLAEPIAIRPTSETMYPYFAKWIRSHRPLIKINQWTSIVRWBF
KHPTPIRTREPLWQEGHTAHTRKEALEMDVILINEYASIVEDLIATPVKGTGSEN
EKFPFGDITKSIEGFIPEIGRAVQAATSHLLQNTFSKMFGEFEDEKNGKEYAHTSW

```

```
GLTTAIGVMIMTHGDNKGLVLPKVPVQVVIIPKTVITBEOKKICNEVECIK
KAGVYKIDRSNTPGKYNHVKVGLCFEPRDIEXRSVVRVVRDNNKMDIP
ISELESIKPLLEEFORNLFPKAKORONESIIRVDTFDKMDTILNOKMVIAPKCEDV
SCBEIKKETARLSDNEDNQSMGKSLCIPNDQIFKIBEGTKCFPCDKLAKFT
LFRGSY"

Alignment Scores:
Pred. No.:      0.00385      Length:      277703
Score:          122.00      Matches:      77
Percent Similarity: 34.7%      Conservative: 51
Best Local Similarity: 20.9%      Mismatches: 107
Query Match:      7.8%      Indels:      134
DB:              13      Gaps:      17

US-10-734-719-9 (1-291) x BX538353 (1-277703)
Qy 25 PheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLys 44
Db 178611 TTTAATATATTTCAACAATTTCAAGATTATGATAAATATTCAAAATATTTATAAAAAAT 178552
Qy 45 CysLysAlaValPheTyrThrPro-----AsnPhePheGluGlnTyr 59
Db 178551 AACAAA---ATCTTTAATACACCTATTCAAGGAATCTTTAATATTATTATAATTTGTTT 178495
Qy 60 Tyr---ThrLeuLysHisLeuileGlnAsnGlnGluTyrGluThrGluLeuileMetCys 78
Db 178494 AATCCCAATTAATAATCTCAAAATACAAATCATGAATATGATCCAGTCTATTTTTC--- 178438
Qy 79 SerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyr----- 96
Db 178437 -----AATATTGAAATATTATAAGAACTCGATTCTATCAAACT 178396
Qy 97 AspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu----- 114
Db 178395 CAATCTCTTTGATAATTAATCAACATAT---TTTTAGAAATATTATAAGAACTGAT 178339
Qy 115 PheAsnAlaTyrPheLysPheHisGluileTyrPheAsnGlnArgIleThrSerGlyVal 134
Db 178338 TTTGATGGTCATTTTCAATA-----TTCAACTTTAAAGATGTTAAATAATATT 178291
Qy 135 TyrMetCysAlaValAlaIleAlaLeuGlyTyr----- 145
Db 178290 -----ATTAGTATTCACCTGAAATATAAATAATAACTAAATCAAGATATTAT 178243
Qy 146 -----Lys 146
Db 178242 ATTATTGAAACTTTTCAATCTTTTAAATAATGAATATTATTCTTATAATAATATGAA 178183
Qy 147 GluileTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyr----- 162
Db 178182 GATCTAATTGTTAAATTTCTTGATCATTTATGATTCAAATTCACAAATTATTATTCAAAT 178123
Qy 162 ----- 162
Db 178122 TGGAGAAATCTAGATTATTATATATAGAAATATATTATTCAAGCTTAGATTCATTA 178063
Qy 163 -----AlaPheAspThrLysGlnGluAsnLeuLysLeu----- 174
Db 178062 TGTAAGCTTTTCTCAAGAACCTTAATCTTATATAAAATGATTAAAGAGAAATAATTT 178003
Qy 174 ----- 174
Db 178002 GTTAATGATGAATTTATATATTAATTTATGACATTTTAACAGAAAAGATATTATAAT 177943
Qy 175 -----AlaProAspPheLysAsnAspArgSerHis 184
Db 177942 TTTAAAGAAAGAAATTTCTATTTCTGAATCTATTGATGATTAATTTCAAGATTGGAATT 177883
Qy 185 TyrileGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeuGlyLysThrTyr 204
Db 177882 ATTATTATTATATATTTCAATAATGAATTAATAGAAATAGAAATTAATTTTCAGAAATTCAT 177823
Qy 205 LysileLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPheIleGluLeuAla 224
```

```
Db 177822 GAAATAAAACTA-----TTAGATCAAAATATTCAAGAAATGATA 177784
Qy 225 ProAsnLeuAsn---SerAsnPheIlelleGlnGluLysAsnAsnTyrThrLysAspIle 243
Db 177783 GAAAAATTGAATGGTGAGAAATTTT-----AGAATAAGAAAGAAATTTGATGAGAAATATT 177730
Qy 244 LeuileProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLysLys--- 262
Db 177729 TTGGATTATGATTAGATTTCACAAATGATTTAAATCTTCAATAATAATAAAAAATCA 177670
Qy 263 -----IleLysIleLysGluAsnValTyr-----LysLeuileLysAsp 276
Db 177669 TTGAATGATATGATTATTTATTCATCAACCTATAAGTTATGATCAAGAAATGATTCGAGAT 177610
Qy 277 LeuLeuArgLeuProSerAspIleLys 285
Db 177609 CAACCTTAATATAGATTTTCTCTATAAAA 177583

RESULT 131
PFMAL4P3 349751 bp DNA linear INV 16-APR-2005
Plasmodium falciparum MAL4P3.
ACCESSION AL035476 AL844503
VERSION AL035476.7 GI:23498195
KEYWORDS
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1
AUTHORS Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,D., Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K., Buckee,C.O., Burrows,C., Cherevach,I., Chillingworth,C., Corton,C., Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corton,C., Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J., Fellwell,T., Goble,A., Goodhead,I., Gwilliam,R., Hamlin,N., Hance,Z., Harper,D., Hauser,H., Hornsby,T., Holtroyd,S., Horrocks,P., Humphray,S., Jagels,K., James,K.D., Johnson,D., Kethnorou,A., Knights,A., Konfortov,B., Kyes,S., Larke,N., Lawson,D., Lennard,N., Line,A., Maddison,M., McLean,J., Mooney,P., Moule,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A., Rabinowitsch,E., Rajadream,M.A., Rutter,S., Rutherford,K.M., Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R., Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,L., Whitehead,S., Woodward,J., Sulston,J.E., Craig,A., Newbold,C. and Barrell,B.G.
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
Nature 419 (6906), 527-531 (2002)
12368867
2 (bases 1 to 349751)
Hamilin,N., Pain,A., Berriman,B., Hall,N., Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M. and Barrell,B.
Direct Submission
Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
On Oct 3, 2002 this sequence version replaced gi:5763808.
For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.
Location/Qualifiers
1. 349751
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4P3"
/complement(join(2130..3551,4472..9988))
/gene="PF0630c"
/feature="synonym: VAR"
/complement(join(2130..3551,4472..9988))
/gene="PF0630c"
/feature="similar to Plasmodium falciparum var, mal4p2.58"

FEATURES
source
1. 349751
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4P3"
/complement(join(2130..3551,4472..9988))
/gene="PF0630c"
/feature="synonym: VAR"
/complement(join(2130..3551,4472..9988))
/gene="PF0630c"
/feature="similar to Plasmodium falciparum var, mal4p2.58"

gene
complement(join(2130..3551,4472..9988))
/gene="PF0630c"

CDS
complement(join(2130..3551,4472..9988))
/gene="PF0630c"
/feature="similar to Plasmodium falciparum var, mal4p2.58"
```

mal4p2.58, var SWALL:Q9U0G5 (EMBL:AL035475) (2277 aa)  
fasta scores: E(): 1.5e-143, 52.77% id in 2486 aa, and to  
Plasmodium falciparum variant-specific surface protein  
var-7 SWALL:Q26034 (EMBL:I42636) (2182 aa) fasta scores:  
E(): 1.9e-126, 46.72% id in 2412 aa  
/codon\_start=1  
/product="erythrocyte membrane protein 1 (pFEMP1)"  
/protein\_id="CAD49166.1"  
/db\_xref="GI:23498196"  
/db\_xref="GOA:Q811V7"  
/db\_xref="InterPro:IPR004258"  
/db\_xref="UniProtKB/TrEMBL:Q811V7"  
/translation="MGPBSTADYSSAKHALDSIGKRVHAQVONEAKQRNSGDLKGLL  
TSATUSGSAFTENPCELIKKREKLLGARGERNPCNGSGASEKEFSVSGECDDK  
KLEGRRNGGACAPVRRLLCNKNFQKNNYDSSAKHNLIVDVMWAANYEAQSL  
DQYATYPSGDSFSCMTMLARSFADIGDIIRGKDLVGRKKKKKQNGKQETETEDQ  
FQIYNELTNGRVGKHQYQDNGGNYFQLRDEWNTANRATVWKAITECKADGNAYFR  
PTCSNRQGSQAHYKRCNGDPDDKPNTPPTFYVQYLRFWEAEADCRKKK  
IYVGIKTYCREKYSKGNPRYCSRNGYDCTTKRAIGKRMGNCICSLACNPVVD  
WINNQEQLDKKKYTDINEASRRQRKRGRSTGSSSYDGESEFYNNLKDDY  
KVDVDFLLNKEKSCQAVKNDGGTINFTEKNDKNNKOKGTFYRSEYCOVCPDC  
GVKYSNGSWEKKKQDQCNILKYPKDAHPHTPIKILKSGEGKEIEKLEAFCKDKKQ  
GNSDSLSLPWCYQDFQLEKDEKEGVDRNDVDTGGGLIKLQKNGEENGKK  
OKTYNDFNFWAHMLKDSIHWKKLEKCLQNGTKRCNNECKNECECFQKRWKK  
KTEWGLKDHFKQNTIGDTCPIVTLGVLKQLFNDSTODKNSLDSBELKHLK  
HUSEMLQETSGDLTCGASNEKETMDKLIHEBEGIAKECLRKQNECEKAPEGRS  
DSHDDPQDDADNEDDLDEDEEVEVQVEDTOEEGEQPVVQOEGSSPTPA  
PAGPVCDDIVSLQSPFSFSDACTILKYVGNKWKCI PSNGNTSDTTGSESEATGAR  
QRDTDSGDTTFLWHKYKAENTTQGVGAGADFLPATSSVATAPGAVPSRSLQ  
AFTQSAIEFTFLWHKYKAENTTQGVGAGADFLPATSSVATAPGAVPSRSLQ  
LSGVGVPFGMAFPVKSI PVPLGVGVGIPGVGALGGAGGPGVGLDVGGOAQ  
PLTLRPGLVGLQSQRLTLDHGHPAGESEDPTQWLOQGHIPDFLRMLFVT  
LADYRDICTGGDRDITVNTSGSSSKIKISEKILHNNDKQEPAPKPSVEK  
TYPEVWQSGDQKNGMI CALTYDPSGISTEDKVRDLIDKDTGKPKQNGDNYT  
YKTELKLDSESGPKGNDTIQPAATLQDFVEIPTVRLHWGSDFCGKRAMLKDVKN  
CENIKDAGHYKSCDGVCTDRVIERNDKFVDLNCGLCYKQCKYKWIIDIKFVEYHN  
QEKYKDEYGLTKDSDDKLGEYKCAENFLKELCHKPSEDDNDKINFDKPE  
KTFNPKSKACAPVGVYIGGNYIPNEEKYKSKGRVKENDTIPNIEVLVLRK  
GEEKDKHLHADKNTGLFEVARYEQWNCQKKGIDQCKITFPANDIDPKDILVFN  
FFORWLRYFVDYKDLKINPCTKETETEKEKSKYCTQGCNDKCEVKEWLSKKQ  
EWQIITLYKYSKISDOEIAFRVYSFVQGLFDNDYKKAQVEVPCDKELMGCT  
GNLKEGEQKCHMGDFITNLISKQIKDDCNKQAOVSVEVTPSDENTACODTH  
PDDEDLLEENENVAQPNICNQVEDKKIEEVEKCEAQTAAETAAAGGEROTIP  
PAPAPAAFPPLPKPKPKPLPPALKNAMLSSTIMWSVGLGFAFTFYFLFKKKT  
KASVGNLFOILPKGDIYDITLKSNNRIPYASDRYKGTIYMEGDSGDEKAYFM  
SDTDDWNLQKDFIISMLQNTQNTBNI LHDVNDNTHPTMSRHNMDQKPFIMSID  
PITDDWNLQKDFIISMLQNTQNTBNI LHDVNDNTHPTMSRHNMDQKPFIMSID  
RNLPSEENYDNFNSNNPINISDTSNMSDLSNNHS PYNKDNDLYSGIDLINDAL  
SGNHIDIEMLRKNEKELFGTHQHPKNTSRNVVTQTSDDDPHNLQNLPHFWLDRH  
RDMCEKLNDRERLAKLEWENETHSGDINSGPSNHNLTNDVSIQIHMDNPKPIN  
EFTNMDTSPDKSTMDTIIDDLKYNEPYDYFYDNDNDKTSMDNNNNLVNK  
NPNVDGNSSTYNHNPADINKTFVDINNHNQHPLEKPTKIQIEMNNREVEQQYPI  
ADIWNI"  
13697..13898  
/note="GC-rich region"  
complement(join(17269..18690,19611..25169))  
/gene="PF00635c"  
/notes="synonym: VAR"  
complement(join(17269..18690,19611..25169))  
/gene="PF00635c"  
/notes="Similar to Plasmodium falciparum var, mal4p2.58  
mal4p2.58, var SWALL:Q9U0G5 (EMBL:AL035475) (2277 aa)  
fasta scores: E(): 1.1e-129, 49.19% id in 2498 aa, and to  
Plasmodium falciparum variant-specific surface protein  
var-7 SWALL:Q26034 (EMBL:I42636) (2182 aa) fasta scores:  
E(): 7.9e-126, 46.25% id in 2441 aa"  
/codon\_start=1  
/product="erythrocyte membrane protein 1 (pFEMP1)"  
/protein\_id="CAD49167.1"  
/db\_xref="GI:23498197"  
/db\_xref="GOA:Q811V6"  
/db\_xref="InterPro:IPR004258"  
/db\_xref="UniProtKB/TrEMBL:Q811V6"

/translation="MAPTSGGGTKDSAKHALDSIGKRVHAQVONEAKQRNSGDLKGL  
LLTSATSGESAFTENPCELIKKREKLLGARGERNPCNGSGASEKEFSVSGECDD  
DKLEGRRNGGACAPVRRLLCNKNFQKNNYDSSAKHNLIVDVMWAANYEAQSL  
IPHQDYATYPSGDSFSCMTMLARSFADIGDIIRGKDLVGRKKKKQNGKQETETEDQ  
LESLKEIFGDIYNELTNGRVGKHQYQDNTDYPQLRDEWNTANRATVWKAITEC  
GDTYFRPTCSNRQGSQAHYKRCNGDPDDKPNTPPTFYVQYLRFWEAEADCRKK  
FYCRKKKYVDIVVKYKRGVYNDVPRYCSRNGYDCEQTIYKGYFYVDIGDINCLVACN  
FYWEIIEKQEPDKQKQYTDINEASRRQRKRGRSTGSSSYDGESEFYNNIL  
KDYQTVDDFLKLNKEKSCQAVKNDGGTINFTEKNDKNNKDKGTFFYRSEYCOV  
CPDGVKYSNGSWEKKKQDQCNILKYPKDAHPHTPIKILKSGEGKEIEKLEAFCK  
DKKGGNSDSLSYPWCYQDFQLEKDEKEGVDRNDVDTGGGLIKLQKNGEENGKK  
EAGNTSEKDEHDEIKQTFNPFYFVWAHMLKDSIHWKKLEKCLQNGTKRCNNECK  
NKECECFQKRWKKKTEWGLKDHFKQNTIGDTCPIVTLGVLKQLFNDSTODKNSL  
QNSCEKKAKEGRSDSHDDPQDDADNEDDLDEDEEVEVQVEDTOEEGEQPVVQO  
VPOBEGSSPTPAPAGPVCDDIVSLQSPFSFSDACTILKYVGNKWKCI PSNGT  
SDTTGSSQSGENLLEAF OSAAIEFTFLWHKYKAENTTQGVGAGADFLPATSSVAT  
LAPAVPSRSLQILSGVGPFGMAFPVKSI PVPLGVGVGIPGVGALGGAGLGP  
GAPVGLDVGGOAQPLTLRPGLVGLQSQRLTLDHGHPAGESEDPTQWLOQGHIP  
GHIPDFLRMLFVT LADYRDICTGGDRDITVNTSGSSSKIKISEKILHNNDKQEP  
LTPDPEKAPKPSVEKTIPEWWSQGHILWNGMICALTYDTSGSSITEDKVRDLKIDK  
DTGKPKQNGDNYTTEKELKDDSGPKGNDTIQPAATLQDFVEIPTVRLHWGSDFC  
CGKRAMLKDVKNCRNIDKAGHYKSCDGVCTDRVIERNDKFVDLNCGLCYKQCKR  
YKWIIDIKFVEYHNQKRYKDEYGLTKDSDDKLGEYKCAENFLKELCHKPSEDD  
NNDKNNKINPDKPKTPNPKSKACPVGVYIGGNYIPNEEKYKSKGRVKEND  
ANDIDFKDIVNFPFQRLWRYFVDYKDLKINPCTKETETEKEKSKYKTCQNGD  
KCECEKWLKSKQEWQITKLYKQYSKISDOEIAFRVYSFVQGLFDNDYKKAQEV  
VPEKCDKELMGCTGDNLKEGEDPKCHMGDFITNLISKQIKDDCNKQAOVSVE  
AFTQSAIEFTFLWHKYKAENTTQGVGAGADFLPATSSVATAPGAVPSRSLQ  
AFTQSAIEFTFLWHKYKAENTTQGVGAGADFLPATSSVATAPGAVPSRSLQ  
GFAFTFYFLFKKTKASVGNLFOILPKGDIYDITLKSNNRIPYASDRYKGTIY  
MEGDSGDEKAYFMSDTDDWNLQKDFIISMLQNTQNTBNI LHDVNDNTHPTMSR  
PITDDWNLQKDFIISMLQNTQNTBNI LHDVNDNTHPTMSRHNMDQKPFIMSID  
RNLPSEENYDNFNSNNPINISDTSNMSDLSNNHS PYNKDNDLYSGIDLINDAL  
SGNHIDIEMLRKNEKELFGTHQHPKNTSRNVVTQTSDDDPHNLQNLPHFWLDRH  
RDMCEKLNDRERLAKLEWENETHSGDINSGPSNHNLTNDVSIQIHMDNPKPIN  
EFTNMDTSPDKSTMDTIIDDLKYNEPYDYFYDNDNDKTSMDNNNNLVNK  
NPNVDGNSSTYNHNPADINKTFVDINNHNQHPLEKPTKIQIEMNNREVEQQYPI  
ADIWNI"  
complement(join(28486..29547,29708..29776))  
/gene="PF00640c"  
/note="synonym: RIF"  
complement(join(28486..29547,29708..29776))  
/gene="PF00640c"  
/note="Similar to Plasmodium falciparum rifin pfcl115w,  
mal3p7.54 SWALL:O97311 (EMBL:AL034559) (371 aa) fasta  
scores: E(): 4.8e-58, 49.73% id in 382 aa, and to  
Plasmodium falciparum rifin pf0035C SWALL:O96113  
(EMBL:AF001367) (375 aa) fasta scores: E(): 1.9e-46,  
43.52% id in 386 aa"  
/codon\_start=1  
/product="RIFIN"  
/protein\_id="CAD49168.1"  
/db\_xref="GI:23498198"  
/db\_xref="InterPro:IPR002858"  
/db\_xref="UniProtKB/TrEMBL:Q811V5"  
/translation="MKVHVISILFSPHNLILLTSQVYNNQKNPYLTPHKNPKSVKT  
CRSLCECLYAPSNYDNDPEMKVKVQOPDRDTSQSHYKHRIQENRQKCEQEKDI  
QKILKDKIEKELTEKALQTEIRSDAIPCTCOKSGYKVEKTCCLKGGLIGVGA  
PSLGLGIGGLVNNKNTPIYFAVFAQKEGIAAGKIASDAARIVTGISILSNF  
EVTHTGSLTANAITLETUKDTILTALKHFHYEGSMVNCVNTDIDDKLICAYGMARGLVQ  
IKSSAPAVIRSVKTLKNADNVAQAQTANETTSGMIKAEUSKIASAGANTYSA  
ITYSVTAVTLVLMVIVLILRYRRKKKKKQLQYIKLKE"  
join(34475..34549,34788..35753)  
/gene="PF00645w"  
/note="synonym: RIF"  
join(34475..34549,34788..35753)  
/note="Signal peptide predicted for mal1p4.07 by SignalP  
2.0 HMM (Signal peptide probability 0.895, signal anchor

probability 0.008) with cleavage site probability 0.389  
between residues 24 and 25

## Alignment Scores:

Pred. No.: 0.00482 Length: 349751  
Score: 122.00 Matches: 75  
Percent Similarity: 34.3% Conservative: 54  
Best Local Similarity: 19.9% Mismatches: 137  
Query Match: 7.8% Indels: 110  
DB: 13 Gaps: 16

US-10-734-719-9 (1-291) x PFMAL4P3 (1-349751)

```
Qy 2 LysLysValIleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeu 21
Db 308754 AGGGATGTTATAAAATAAAATGGTGTATTAAATAAAAGGATGCCGTTAATATT 308813

Qy 22 ProAsnAspPheAspValPheArg---CysAsnGlnPhe-----TyrPheGlu 36
Db 308814 ACAATGATGTTAATATTATTTTAATGATGTTAATATTATTAATGATGTTAATATTAT 308873

Qy 37 AspIysTyrTyrLeuGlyLysCysLysAlaValPheTyrThrProAsnPhePhe--- 55
Db 308874 GAT-----GTTAATATTATTTTAATAATGTTAATATTATTTTAAT 308909

Qy 56 -----PheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsn 68
Db 308910 AATGTTGATGAGTCGATGTTATCTCTAATATTATTC-----GATATCATAGTGAA 308960

Qy 69 GlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsn 88
Db 308961 GATATTTTAAAGTTAAAGTTAAATTCATCATATATTATATTATGAGTTTACGAAATAAA 309020

Qy 89 GluAsnPheValLysThrPheTyr-----AspTyrPhe 99
Db 309021 GAGAACATATTATATAGCTTATATTATTGATTGACAAGATGTATATAGAGATAAATAT 309080

Qy 100 ProAspAlaHisLeuGly-----TyrAspPhePheLysGlnLeu 112
Db 309081 TTACATAAACATATAAAGAATTTAATTAATATAAAATATATACGTATATAAACGAAATG 309140

Qy 113 LysGluPheAsnAlaTyrPheLysPheHisGluIleTyr----- 125
Db 309141 AATAGTTTAAATGATTATGATATGATATATTCATATATTTTGTATTATTAATAATATATG 309200

Qy 126 ---PheAsnGlnArgIleThrSerGlyValTyr-----MetCysAlaVal 139
Db 309201 AATTTTAAATAAATGATAAGAAATTTATTATATTATTTTGTATGATAACATATGTAAGATA 309260

Qy 140 AlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGly 159
Db 309261 AAATTAATAATATTATATTGACAGTCCTCTTA-----TTTTGTTTTTAAAGAATCTC 309314

Qy 160 SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLys 179
Db 309315 GAAAAAGTAGTTATGATTGCTTATACAAAAATATTATAT-----GATGAAAG 309365

Qy 180 AsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPhe 199
Db 309366 AATATTTTAAATAAATGAAATCGATAAGAAAAATAAATAAATGATAAATAATATA----- 309419

Qy 200 LeuGluLysThrTyrLysIleLysLeuTyr----- 209
Db 309420 -----TATAAGTGAATTTATATGTAACCATATAAATATTATTATTATTTTAA 309467

Qy 210 -----CysLeuCysProAsnSerLeuLeuAlaAsnPheIle----- 221
Db 309468 ATATTGAATTCATGTAATATTATTTATTTATTTATTTATTCCTACTTTTATTATGAAATTC 309527

Qy 221 ----- 221
Db 309528 ATTTATTCCTTTTAAAAAATATATAAGAAATATAATGATATATTTTATTTTCGACGGTATG 309587
```

```
Qy 222 ---GluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240
Db 309588 AAAAAGTATGTACTACATCAAAATACATCAACATTACAAAGAAAAAAGAACGACG 309647

Qy 241 LysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260
Db 309648 GAAATATATGACCATGAGGAATAATAAAATGAAATAATATGGCGAAAAAATATATAGCGAA 309707

Qy 261 LysLysIleLys-----IleLysGluAsnValTyrTyrLysLeuIle----- 274
Db 309708 AAAATATATATCCAGGATATTAAATAATATATATGCAATAAGAAATATTTTTCATCATCC 309767

Qy 275 ---LysAspLeuLeuArgLeuProSerAspIleLysHisTyrPheLys 289
Db 309768 CATGATGATATATTGAAAAATCAATATGATATGAAAGGTTGTATATAA 309815
```

## RESULT 132

## PFMAL4P2/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

## gene

## CDS

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence

## sequence



```

Plasmodium falciparum reticulocyte binding protein 2
homolog a SWALL:Q9BK46 (EMBL:AF312916) (3130 aa) fasta
scores: E(): 3.4e-23, 22.92% id in 2753 aa, and to
Plasmodium falciparum reticulocyte binding protein 2
homolog b rbp SWALL:AA02259 (EMBL:AF312917) (3254 aa)
fasta scores: E(): 5.8e-21, 21.63% id in 2893 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB62842.1"
/db_xref="GI:6562703"
/db_xref="UniProtKB/TrEMBL:Q9U0M1"
/translation="MDQKNILICGKNDISKETETKEHANVLGCAKTEKEEKNLVDD
KYNEKSNLININNKNDIMKTDDVNDKTLKDDTLIERINQKVTKKQHLINEEKN
IHEIDNQNGKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNK
SKPDNDKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNKSPDDNK
ENIRKQKFMVYKQVYIHYAEKVMNIKEIKRYKEBYENHYFAKSYLCHDDLLNIKDOT
IMFVKHRELAVENTNLCEHKNKDNKI FNEINGRLSILSEDTNMEIKDDEVYKII E
QTHQCQKQENVPVDIMNIHKDTYNDNATNINKEDDNNYLDMDIHKEMKSNVGL
SYIYYMNNNDMFKNVYIIMKSEGNVYTRHMKKIDDEORCIYNTCSKNYISINHY
SGDINTCLGNVSGEDHTFVNPYDKKKWEIRKSTNNYNNINRKNRTHRSUKNR
IINIKHNISYVICNRYLKGGKYERMLKNNLKYNIILKIINKILSCNNIYKYAKM
KYILOSVFNILNIMVKKRTYVIFINTKYKLFINHYKNI SPKFNLMARHFPVADYIF
NKEKPDNDLINVKKKRTYVIFINTKYKLFINHYKNI SPKFNLMARHFPVADYIF
KLSYVITLKEIDLCLYMDNALEILDKNLDIHKENKLIQNHLCTNLEIENI
KQULFHEYSIHRSLFKSIHWEETNKYIIFLNQKKENTQNDHISRLEVLINOSIKIE
EQMDIHKYMYTHYTHYIYSFAYGAKIQOQYIIDEENKQNTERLIIDLEQNHQET
REIMFTEQEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEK
KNIINTEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEK
TNSNDKITLEKEIKNPNDSKITLEKEIKNPNDSKITLEKEIKNPNDSKITLEKEIQNI
RNEKITIEKQISNDKVITIEKQINRNDKMTLEKEIKNPNDSKITLEKEIQNF"
5644. .6063
/gene="PF00325w"
/gene="PF00325w"
5644. .6063
/gene="PF00325w"
/note="Similar to Plasmodium falciparum very hypothetical
protein, Mal4p2.02 SWALL:Q9U0M0 (EMBL:AL035475)
(139 aa) fasta scores: E(): 1.4e-47, 98.57% id in 140 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB62843.1"
/db_xref="GI:6562704"
/db_xref="UniProtKB/TrEMBL:Q9U0M0"
/translation="MYRNKIYFLTNWKRTRYHSRQNTNNINSNKIKDPYSDLYKSS
MYGNFKILNKTSTKREYIEIKTSNNTSYSTPDPNINNTLTPPESSQMYENR
KYIMKYKNVYIPIKRYTKYAKSKTNNTVYIRMEK"
join(6794. .6857,6942. .7075,7272. .7383,7566. .7663,
7792. .8064,8151. .8357,8454. .8588,8638. .9609,9740. .10018,
10170. .10250,10340. .10378,10551. .10766)
/gene="PF00330w"
join(6794. .6857,6942. .7075,7272. .7383,7566. .7663,
7792. .8064,8151. .8357,8454. .8588,8638. .9609,9740. .10018,
10170. .10250,10340. .10378,10551. .10766)
/gene="PF00330w"
/note="There is slight modification in the gene model in
the last exon from submitted (SWALL:Q9U0L9
(EMBL:AL035475)) Plasmodium falciparum hypothetical 103.2
kDa protein sequence
Similar to Plasmodium falciparum hypothetical 103.2 kDa
protein mal4p2.03 SWALL:Q9U0L9 (EMBL:AL035475) (853 aa)
fasta scores: E(): 0, 100% id in 850 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49159.1"
/db_xref="GI:23498188"
/db_xref="UniProtKB/TrEMBL:Q811W4"
/translation="MNVDEKRYKIEKYEKQNDILKXALIEYKDKLEKELKKNDI
LSNDKQNLQNLTLTSLBEQKKSGWNLMLLTNSRENTHKSVAFEELEKKEIKE
NETHKNI EDLHOKSYLKELELLKKDHI EKNKELETITGLNDLVNSSEKHOIKAD
EYVEELIKLIEENKQESQI KKEESLDVYRNTSTWLHNKYPKYMINFDQI PFYOA
YDKGSYIQEILQNERLANYSLQKVVVYLKVCHELFLFOGKILNDSYMNQKKE
KLBEQDKSKMIDUKNICDQOKILBEEIDTAEKFYIYAEKDEKLEMTVFLNINI
QMYTHKYNIFHIIKNKIRLELVKVNVLICIEEYMFPSFINNKVFLNRNVIQQPFLKKE

```

```

LVIKIINIITLIISICPNNERIILENLKRNIIKKENNKIINHMEKRNDEEKTYNIS
SPDMHITNIVYDDDLKLVYKNQENKSYLNNFFIVEKIPFINFQBEKEKIMLSQFND
FTNYQTNHNIILPILKKIOLIOYIQSLKFLKSYISYRICEINFRGNYTNNSPLI
TNLUDDKTFEIKSENFYIHIHNIITKLSYINHTPYLYQHWYTIINNQTENKA
KVHYDGLKNLKDQLQPKDNIQNNMKLKSKEKKNLKTTLNNFRILLNEBLSL
NNQYLSINDKREDKQYSDNNNIQTYAEKIYNFISHTDEGKNFNKEKQILISAYI
SSCIKINNLEIQNNKHILEQLNFKINEIKNLNTQIDTYKEEVEYNHKKYEEQV
MNTLHDLIVSLEKQINKINSEKNVKKPFIMCSICSNKKNIGRKYNPCTCNMLYCVVCV
MCAHFYHFFPHF"
complement(10868. .19509)
/gene="PF00335c"
complement(join(10868. .10972,11088. .11177,11367. .11617,
11727. .11807,11991. .12324))
/gene="PF00335c"
/note="Similar to Plasmodium falciparum hypothetical 33.9
kDa protein mal4p2.04 SWALL:Q9U0L8 (EMBL:AL035475) (286
aa) fasta scores: E(): 4.9e-117, 98.25% id in 286 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49160.1"
/db_xref="GI:23498189"
/db_xref="UniProtKB/TrEMBL:Q811W3"
/translation="MFFNKLVSRTSYELDYTKRYNLVYKHLKYLQWMDYLKREMINY
FNNRKVNTVDIEAKWEKFKKELIQANGYIPAIIMKYGIEERICIPKHLIDEXAF
EEDGHLSLFLSARLFRVHIHGKVVNCVSVHVEADPIEKHIYFLKFAFVVENQITFIN
VPCSIIGLGPAYINGYVOLAMNSIKCKVLGSI PGPFQIDVSKLTVKETPNISKL
KLDLYLDDGNNVIFSEYNDELTEIWTYFPGKIMETPLPDDYDPNPFVNRGKRIQ
LTYKDYWPKQ"
complement(join(12996. .13021,13108. .13198,13312. .13425,
13569. .18095,18336. .18767,18885. .18910,19003. .19037,
19326. .19336,19432. .19509))
/gene="PF00335c"
/note="Alternative start site possible; gene model
slightly modified from submitted version (SWALL:Q9U0L7
(EMBL:AL035475)) Plasmodium falciparum hypothetical 205.8
kDa protein sequence
Similar to Plasmodium falciparum hypothetical 205.8 kDa
protein mal4p2.05 SWALL:Q9U0L7 (EMBL:AL035475) (1734 aa)
fasta scores: E(): 0, 99.25% id in 1742 aa"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD49161.1"
/db_xref="GI:23498190"
/db_xref="UniProtKB/TrEMBL:Q811W2"
/translation="MNFFLIILFFFLMFLKIAMGVHIFGMCIDCSPLFENFRIS
EDRKTVCNKESEKIDNLDIYWNDDICTETLEEDTDESKNGNFNKQDLGITYHNKN
KVNLEKFKEDDYKLSQFYVEKYNNDKNIYINFDNSEGVNNISFDNDNENFYFNEVD
KNKFINFEIYVYKLGKEYITPGNFVHLYTNQKGNLIEEKDIKLKVMRSRKNKKY
KYCNVDYSYNSINNKORNNLNVNKLKRKTKYDKNEKVMFHKRFPSSYNMSARKEK
ALSKAKAREKLWKRKKREKCEISLNVKMEINLCVTKSENIYNIIVENNLK
NEKLEKIQMVHNDTSYLQKUREKKNDMTINKIIEIRNKRMDRIYNDKNDKDKVQI
NVSSLHVLKQSNILKDSKEFLCSNNEIPVNNREKCEIKKEKFTLFGSNTNNTDLYK
DOIENKDSILKNNIEINSEGVNIIKKHITNRSYSDNIKCSDDYKMNNTQVNDYI
WYVYNSQCNKEKREYSYNNNDYNSQSCNKEKREYSYNNNDYNNYNNIVE
YANNVSHYSYNNKEGNILPYDNIITEENIIDNKENDDLTONMEOLQSYKEYKNES
NEDDKTNFTNLHNSIKKEIKIKQDINVDNMEILNFEENKKNNTYSHYFVEKK
NPSSKNEVDNRIRKNKNDMQIINIDIPSNILPTMLLIQHEEKTNNKNDKDKFKYK
NYKVELIGPLNNCVNLKKAIEKELFSRKEIYHVCNMNKKNNICFRKNKMTYK
NNIVSDDEKYGYPVLSRSYIPFPMGDKNNRIHKYKMMNNKYLINFCRDKKDDIYN
NNGREEKNVPLPIKPLNI INVKANEVNTKNYRINKCIYKQCNCDKNILSGIYKD
KALNMKGISKVLKNNIMVSEKQSIKKDNLIKYKNQYFSLHKNNCIKKTSMLSLC
NMKDSQNVCKRNNIIVVHVNKAQMKKEASIKKYNINHNHNIGINRSTHVS
DYKVDYSIQNGLSIDKHDNRVILKSVLFMDNTHGEGMYTIEKFLQGHENSERFIE
KKGREGSFTCEIKNNDQVKNYEMKAGKNDLQNEIIEKEFDLOHENSERFIE
BEKQICIVDDKNNIMNVDEKRSKDPHSYERVLKMGESNKEEGSNTDKILKNEKE
KVNNEKGEDEKNEKNEKNEKNEKDEKNEKDEKNEKDEKNNNNNSFYNNNS
KTNFELCTNSLIFINNKNLSILPSENEKIGSQKEEQNI SPVILNNKDKDLCNNIN
ESDYSQYQSVLLNSIEKKIYKCSNKSIRKEIEKKKINEDIVDLKNINCSNTLFF
LTKYLSKSELIIHEHQCNINNVYKCKKKKQOAKKLNKRNKINVPDSII EENWSS
EYNFVKNNNNKMWKFTKRSKILSSSEIFAVKQKRAITNLMSRQFISSTIGLVEK
GENKRIEKEKBEVYIKKIKNKNKFNKNLTQLLPFKGAENINTSGSFNTEKIRHVK
RTKRKNLSNNFLNNFNNILKQORMEEDKIKMDQKQKEINKKNEKNEKGEKEDIK

```

## Alignment Scores:

```

Pred. No.:      0.00379      Length:      234081
Score:          121.50      Matches:       71
Percent Similarity: 34.9%      Conservative: 59
Best Local Similarity: 19.0%      Mismatches: 100
Query Match:      7.8%      Indels:    143
DB:              13      Gaps:      19

US-10-734-719-9 (1-291) x PFMAL4P2 (1-234081)

QY 30 CysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysLysCysLysAlaValPhe 49
Db 90368 TGTAAATCCTCTTAT-----AAATTTTAATATTTTGTCAATGTAAT 90330
QY 50 TyrThrProAsnPhePheGluGlnTyr----- 59
Db 90329 AATAACCAATAATATTTTAAAACTTTTCTTATAGACGACAAAGAAAATGTTATATTA 90270
QY 60 -----TyrThrLeuLys----- 63
Db 90269 ACAAAATGAAAATTTACATATACTCTCAAGCGTTCCTTGAGAAATTTTCATCCAGTATGT 90210
QY 64 -----HisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIle----- 76
Db 90209 TATAAAAATAAATCACATTTATCTAGAAAATGAACAAACATGAAGGAAAATTTATATAAT 90150
QY 77 -----MetCysSerAsnTyrAsnGln----- 83
Db 90149 ACAATAATAGAGTATTTAATAATCATACATTCGCAAAATATTAATAAAGAACTTAAT 90090
QY 84 -----AlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPheProAsp 101
Db 90089 AAATATAGTATCTACTAACAACTCAACATATT-----AATATACATTAATTTATTCCTCC 90039
QY 102 AlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPhe 121
Db 90038 -----TCATCTTATTTCTGTCGTTAATAGTTGAATATAA--TATATAAAGAT 89991
QY 122 HisGluIleTyrPheAsnGlnArg-----Ile 130
Db 89990 GAAGAAACTTATATATGAAAGGAATATTTTACATTTTATGATCAAAATATATATA 89931
QY 131 ThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGluIleTyrLeu 150
Db 89930 AATAGTTTCACACATTTATATTTGTTTATTTCTCTTGTAGTTTATATAAGATATATGTA 89871
QY 151 SerGlyIle-----AspPheTyrGlnAsnGly----- 159
Db 89870 TGTAAATACATAAGATTTTATTTTGTATTTATTTTAAAGGCAAAATATTAATCA 89811
QY 160 -----SerSerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLys 173
Db 89810 ATAAATGTAATATTTAAACAATGAATATCAATGTGAAGATGATAAAGATAGATTTCTTCAA 89751
QY 174 LeuAlaProAspPhe---LysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThr 192
Db 89750 AATATACATCATATATGAAAAAATTTGAACAGACATATATTT----- 89709
QY 193 AspIleLysAlaLeuGluPheLeuGluLysThr-----TyrLysIleLysLeuTyr 209
Db 89708 ---ATTAAAGAAATGATATGTCNAAAAAAGGATACATCTTTTAAATAATTAATTTGTTCT 89652
QY 210 CysLeuCysProAsnSerLeuLeuAlaAsn----- 219
Db 89651 TCTTTATATATTAATTTCTACATATATAAATGAAGGTATTTATATATATATAATAATAAT 89592
QY 220 -----PheIleGluLeuAlaPro-----AsnLeuAsnSerAsnPheIleIleGln 234
Db 89591 AAACATTTTTTTGAAATCAAAACCCTCTCATTTAAATATATCTCTATAATAATCAACAC 89532
QY 235 GluLysAsnAsnTyrThr-----LysAspIleLeuIleProSerSer 248
Db 89531 ATAAAAATTTCTCTGTCCTCTTCATTTTATTAATATGAAGAATTTACTTATTTTAAAGACTA 89472
```

```

QY 249 GluAlaTyrGlyLysPheSerLysAsnIle----- 258
Db 89471 GAATTTTTTAAATAATAATCAAAAATTTGTTAATCAAAATGCTTTACAACTTTTCAA 89412
QY 259 -----AsnPheLysLysIleLysIleLys-----Glu 267
Db 89411 ATAAAAATTTATATATATATACAAAATTAATTTAAATCTTATACAAAATGATTTTAAAG 89352

QY 268 AsnValTyrLysLeuIleLysAspLeuLeuArgLeu 280
Db 89351 AACTCTAATTAATAATATATCAAAAATATATAAATCTTTA 89313

RESULT 133
AE014851/c 251762 bp DNA linear INV 11-FEB-2003
LOCUS Plasmodium falciparum 3D7 chromosome 12, section 8 of 9 of the
DEFINITION complete sequence.
ACCESSION AE014851 AE014188
VERSION AE014851.1 GI:23496930
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 251762)
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Fraser,C.M. and
Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
PUBMED 12368864
REFERENCE 2 (bases 1 to 251762)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 3 (bases 1 to 251762)
AUTHORS Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES
    source
        1..251762
            Location/Qualifiers
                /organism="Plasmodium falciparum 3D7"
                /mol_type="genomic DNA"
                /isolate="3D7"
                /db_xref="taxon:36329"
                /chromosome="12"
                /complement(70..161)
                /rpt_type=tandem
                /rpt_unit_seq="(TA)n"
                /note="AT_rich"
                /complement(206..243)
                /note="AT_rich"
                /rpt_type=tandem
                /rpt_type=tandem
                /complement(262..350)
                /note="AT_rich"
                /rpt_type=tandem
                /rpt_type=tandem
                /complement(355..391)
                /rpt_type=tandem
                /rpt_unit_seq="(A)n"
                /complement(424..453)
                /note="AT_rich"
    repeat_region
    repeat_region
    repeat_region
    repeat_region
    repeat_region
    repeat_region
```











```
BA000026_06 600001 710000
BA000026_07 700001 810000
BA000026_08 800001 910000
BA000026_09 900001 1010000
BA000026_10 100001 1110000
BA000026_11 1100001 1210000
BA000026_12 1200001 1310000
BA000026_13 1300001 1358633
Continuation (12 of 14) of BA000026 from base 1100001 (BA000026 Mycoplasma penetrans HF-

Alignment Scores:
Pred. No.: 0.00245 Length: 110000
Score: 120.50 Matches: 62
Percent Similarity: 39.1% Conservative: 40
Best Local Similarity: 23.8% Mismatches: 78
Query Match: 7.7% Indels: 81
DB: 15 Gaps: 14

US-10-734-719-9 (1-291) x BA000026_11 (1-110000)
QY 51 ThrProAsnPhePheGluGlnTyrThrLeuLysHisLeuLeuGlnGlu 70
Db 48139 ACATTAGACTTCAGTTCAATAAGAACTAACTACAAGAAATTTTAGCGCTTGAAAAA 48080
QY 71 TyrGluThrGluLeuLeuMetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsn 90
Db 48079 GAAGTTAAAGAAATTTGCTACTAAAAATCCAATCAAGTT----- 48038
QY 91 PheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAspPhePheLys 110
Db 48037 TTAATGAAGAACT----- 48020
QY 111 GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgile 130
Db 48019 GAAGCAAAAGAAATGGGTCTATTGTTTATTTGAACAAGTTAT-----AAAAA 47969
QY 131 ThrSerGly-----ValTyrMetCysAlaValAlaLeuLeuLeuLeuLeuLeuLeu 145
Db 47968 ATTAGTGGGAACCTAAGAGTTTATATTATGTCAGAACTATTGAAATATGTGGTGA 47909
QY 146 LysGluLeuTyrLeuSerGly-----LleAspPheTyrGlnAsnGly----- 159
Db 47908 ACTCATGTTTACATACTGGAGATATTGAAGACTTTATGGTTACAGGACTAATCTCTAAG 47849
QY 160 -----SerSerTyrAlaPheAspThrLysGln 168
Db 47848 GGTCTGCTTCTTGAAGAATTGAACAGCTTCTTCAAATTTATTAGTAGATAAATTAAAG 47789
QY 169 GluAsnLeuLysLeuAla----ProAspPheLysAsnAspArgSerHisTyr----- 185
Db 47788 AATAATGTAATCAAAAGCAATTCATGATGATTTTAATAATTTCTTAAATAATATAAGAA 47729
QY 186 -----LleGlyHisSerLysAsnThrAspIleLysAlaLeuGlu 198
Db 47728 TTAATATTAAGATGATGAAGTAGAGAAAAATATAAAGACAGATTTAATTTCTATCAT 47669
QY 199 PheLeuGluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAla 218
Db 47668 TATTTAGAA-----CTTAAAGAAATT-----AATGAAATTTTAAAA 47633
QY 219 AsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsn 238
Db 47632 AAT-----AAATTAACACTTTTAGTGATTTAGAAAGAGAGAGAGAGAGAGAGAG 47591
QY 239 TyrThrLysAspIleLeuLeuProSerSerGluAlaTyrGlyLysPheSerLysAsnIle 258
Db 47590 TTAAGTAAAGAA-----TCTAATGAAATTAATAAATAAATTTACT----- 47552
QY 259 AsnPheLysLysIleLysIleLysGluAsnValTyrTyrLysLeuLeuLysAspLeuLeu 278
Db 47551 -----GAAGTAAAGAAATCAACAAAACTTTCTTATTAAAGATATTGAT 47507
QY 279 Arg 279

Db 47506 AGA 47504
RESULT 138
AC006278/c
LOCUS
DEFINITION
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 4 unordered pieces.
AC006278
AC006278.8 GI:9797715
HTG; HTGS_PHASE1.
Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 153477)
Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
Kurt, O.B., Conway, A.B. and Davis, R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
TITLE
JOURNAL
AUTHORS
REFERENCE
2 (bases 1 to 153477)
Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurt, O.B. and Davis, R.W.
Direct Submission
Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT
On Aug 12, 2000 this sequence version replaced gi:8810448.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10515: contig of 10515 bp in length
* 10516 10715: gap of unknown length
* 10716 144870: contig of 134155 bp in length
* 144871 145070: gap of unknown length
* 145071 149010: contig of 3940 bp in length
* 149011 149210: gap of unknown length
* 149211 153477: contig of 4267 bp in length.
FEATURES
Location/Qualifiers
1..153477
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clone="PFYAC341"
/clone="3D7"
10516..10715
/estimated_length=unknown
144871..145070
/estimated_length=unknown
149011..149210
/estimated_length=unknown
ORIGIN
Alignment Scores:
Pred. No.: 0.00338 Length: 153477
Score: 120.50 Matches: 67
Percent Similarity: 36.6% Conservative: 39
Best Local Similarity: 23.1% Mismatches: 103
Query Match: 7.7% Indels: 81
DB: 12 Gaps: 13
US-10-734-719-9 (1-291) x AC006278 (1-153477)
QY 21 LeuProAsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr 40
Db 140823 TTGATTAATGAAGAGAACATTTTATATATAAATATTTTATATGAAAAAATATATCAT 140764
QY 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
Db 41 LeuGlyLysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyr 60
```



Db	140763	ATAATTGAAATAATAATTACTTTTTTAAAT	-----	140731
Qy	61	ThrLeuLysHisLeuLeuGlnAsnGlnGluTyrGluThrGluLeuLeuMetCysSerAsn	80	
Db	140730	AAATTAATAAATATGATGGGTATGAGAAATAAGTTTCTTAATAATAAATATGT	-----	140677
Qy	81	TyrAsnGlnAlaHisLeuGluAsn	-----GluAsnPheValLys	93
Db	140676	-----ATACACTTAGAAATAATATATATGATAAAATCTTGACAGAAATAATAGATAAA	-----	140626
Qy	94	Thr-Phe	-----TyrAspTyrPheProAspAlaHisLeuGlyTyr	106
Db	140625	CATATTGATATGAGAAATATGAAACGTATATTTTAAAGATGATGAAAGAAAGGACAA	-----	140566
Qy	107	AspPhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPhe	126	
Db	140565	TTATTTTATAAATTTGGTTGTA	-----TATTTT	140536
Qy	127	AsnGlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLys	146	
Db	140535	AAT	-----ATTATAATAAGAAAAAATAAAGAG	140509
Qy	147	Glu	-----IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
Db	140508	GAGTTTATCTTTTGGTTAGTATATATTTATATCATATATCATTTTAAAG	-----	140458
Qy	161	SerTyrAlaPheAspThrLysGlnGluAsnLeuLysLeuAlaProAspPheLysAsn	180	
Db	140457	---TATTATTTTAGACACAAAAATATTAATATAGAAAAAATAAATAAATTTGCTTACA	140401	
Qy	181	Asp	-----ArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeu	197
Db	140400	CAAGTTTAAAAAAATTCATTA	-----	140377
Qy	198	GluPheLeuGluLysThrTyrLysIleLysLeuTyrCysLysCysProAsnSerLysLeu	217	
Db	140376	AATATAATAAGATATATTTAATTAATTTATGTTATGTAATAAATTTATGTAATTTGTA	140317	
Qy	218	AlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGlyLysAsn	237	
Db	140316	ACTAATAATAATAAATTTGGAGCCCTATTTTAAATCAACAGTTTATATAGATGATATGAAT	140257	
Qy	238	-----AsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTyrGly	-----	252
Db	140256	ATAGACCATATACATAAGAATATATATAAATAGATGTTCTTCTATTTATCTTTTCAGAAAA	140197	
Qy	253	LysPheSerLysAsnIleAsnPheLysLysLysLysLysGluAsnVal	-----Tyr	270
Db	140196	AAAAAACAGAAATTAATATTTTAGAAGAAATAATATCAGTCATATGTCAGTTGCTTT	140137	
Qy	271	TyrLysLeuIleLysAspLeuLeuArgLeu	280	
Db	140136	AAAAACTTTATAAATATGTAATACGTATA	140107	
RESULT	139			
PFMAL3P2/c		158548 bp	DNA	linear
LOCUS				INV 16-APR-2005
DEFINITION				Plasmodium falciparum MAL3P2, complete sequence.
ACCESSION				AL034558 AL008989 AL009007 AL009008 AL009009 AL009014
				AL010140 AL010150 AL010152 AL010156 AL010158 AL010211 AL022217
				AL844502 Z97349 Z97350 Z98546
VERSION				AL034558.3 GI:15375364
KEYWORDS				HTG.
SOURCE				Plasmodium falciparum 3D7
ORGANISM				Plasmodium falciparum 3D7
REFERENCE				1 (bases 1 to 158548)
AUTHORS				Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
				Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T.,
				Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Felwell, T.,
				Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S.,
				Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J.,
				Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A.,
				Rajandream, M.A., Rutter, S., Skelton, J., Squares, R., Squares, S.,
				Stulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and
				Barrell, B.G.
				The complete nucleotide sequence of chromosome 3 of Plasmodium
				falciparum
				Nature 400 (6744), 532-538 (1999)
				10448855
				2
				Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,
				Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,
				Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Corton, C.,
				Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,
				Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,
				Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,
				Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,
				Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D.,
				Kerhornou, A., Knighes, A., Konfortov, B., Kyes, S., Larke, N.,
				Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,
				Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A.,
				Rabinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M.,
				Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,
				Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,
				Whitehead, S., Woodward, J., Stulston, J.E., Craig, A., Newbold, C. and
				Barrell, B.G.
				Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
				Nature 419 (6906), 527-531 (2002)
				12368867
				3 (bases 1 to 158548)
				Lawson, D., Bowman, S. and Barrell, B.
				Direct Submission
				Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,
				The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
				CB10 1SA, UKXX
				On or before Aug 30, 2001 this sequence version replaced
				GI:2982545, GI:2982529, GI:2982547, GI:2982546, GI:2982548,
				GI:2624372, GI:2982553, GI:2982560, GI:2894487, GI:2665315,
				GI:2982565, GI:2982575, GI:2982578, GI:2982517, GI:2982518,
				GI:2982519, GI:4493878.
				For more information about this sequence or the Malaria Project,
				see <a href="http://www.sanger.ac.uk/Projects/P_falciparum">http://www.sanger.ac.uk/Projects/P_falciparum</a> .
				Location/Qualifiers
				1..158548
				/organism="Plasmodium falciparum 3D7"
				/mol_type="genomic DNA"
				/isolate="3D7"
				/db_xref="taxon:36329"
				/chromosome="3"
				/clone="MAL3P2"
				join(1030..2337,2429..2502,2645..2685,2826..2941,
				3046..3173,3308..3388,3508..3571,3699..3737)
				/gene="MAL3P2.1"
				/note="synonym: PFC0160w"
				join(1030..2337,2429..2502,2645..2685,2826..2941,
				3046..3173,3308..3388,3508..3571,3699..3737)
				/gene="MAL3P2.1"
				/note="Revised: revised to 8 exon structure, revised:
				donor for exon 7
				4 weak Pfam matches to entry PF00023 ank, Ank repeat; Pfam
				match to entry PF01529 zf-DHHC, DHHC zinc finger domain,
				score 66.70, E-value 4.8e-16
				Similar to Plasmodium vivax pvlh14090_p pvlh14090W
				SWALL:AAF99464 (EMBL:AY003872) (611 aa) fasta scores: E():
				5.6e-172, 71.94% id in 613 aa, and to Mus musculus
				5033406114rik protein 5033406114rik SWALL:Q9DJQ8
				(EMBL:AK017158) (368 aa) fasta scores: E(): 1.7e-09,
				24.87% id in 197 aa, and to Rattus norvegicus small rec
				srec SWALL:Q9JKR5 (EMBL:AF228917) (366 aa) fasta scores:
				E(): 8.4e-09, 27.95% id in 161 aa"
				/codon_start=1
				/product="binding protein, putative"
				/protein_id="CAB38988.3"
				/db_xref="GI:23477021"
				/db_xref="COA:O97224"
				/db_xref="InterPro:IPR001594"

```
/db_xref="InterPro:IPR002110"
/db_xref="UniProtKB/TrEMBL:O97224"
/translation="MNWNSLDSEVLKQYEIIKYAKQDFRLQILQIOPVLLNNDIE
MLNSINILHWACYCGFTLVKLIISPCNDIEKEDLVNNDTPYYAIAKNSVEIIVLLII
KCGSLIFLHNRMRSPFLTAICEFNEDKILEAHLELLIMNGVSLBQNEHQGTA
LFLSVKNNILQWLTUKEVINHRDFTYGNVTLVIAVRHCDIDILRLQDGLNMY
YSSIEKNTNFVQICIKRFLVLYLLKQWLQKISGLKIKTIYAFYFWPFALL
NLIVFNAHSFSLINKYHKSLIWTITWFOQLCMWLFKSPGCFYKAFENMKNKK
NIOSSKYKNSPMYNGTETWAEVQLNRIEREIPOINKLLISTNENTLTPINHDOALS
NKYNDLILNLEYQKLSYQSVQSRINSLDEYRNAILYNQNPRNICVTNIIKPRV
HHCACFHCIVQDHHCVNDNCIGIKNORCFYMFICIFVLLLLNYNYVYLPHLFO
ALINAFGSLVLCNFINVTLFAITFLYFAIRNTILTNVTFYEHKPNHIDKYNT
DTRCWDFONLMLGLKILNYVFWSLNYDEPYIRQCTKTSNDVCYTLISNT"
5341..12021
/gene="MAL3P2.2"
/notes="synonym: PFC0165w"
5341..12021
/gene="MAL3P2.2"
/notes="Revised: original entry split into PFC0165w and
PFC0166w
Similar to Plasmodium vivax pvlh14095_p pvlh14095W
SWALL:Q962L5 (EMBL:AY003872) blast scores: E(): 0.0,
score: 2249 43% id, and to Plasmodium falciparum
thioredoxin-like redox-active protein fred SMALL:Q9NC62
(EMBL:AF234633) (179 aa) fasta scores: E(): 7.5e-47,
99.42% id in 175 aa"
/codon_start=1
/product="hypothetical protein, conserved"
/protein_id="CAB38989.3"
/db_xref="GI:23477022"
/db_xref="GOA:O97225"
/db_xref="InterPro:IPR007259"
/db_xref="UniProtKB/TrEMBL:O97225"
/translation="WMREKLIKLEKENKRMIAIKVFRDCHGMNNYSQCYVNVK
NKNYNDPFWYENIYGDNIITKRSVQEEEEKKMIEDNKNKESSELEFYTPYRYSNRN
IFLINSINENIYGVTSVYNNDLIYIHNKRYEHMNLPKRTKKKTKKTKTKETKR
KKTLLHKNIDTFGSVSNDEKDYKNQDYDKDDDDYDDDDDDDDDDDDDDDDDD
YEDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD
KRKKEKKNISIDKRNIIYSDSDSDNNYEHIFTHNSDLFLSDISNDEKENNLIQ
NEQFINDYDVIKRFKFKNLVNDSCNSWIVIKKSILINEYITRSQNDLVDINDV
LYDVNNMYHRIYNGKLYIINDDFLINNTKEILGVQNVASHNNTNNNNINDI
HNVINNYHNRCHLRIEKKKNNIYNNNNNNNDDELSTDTGDDMLSNCEYNE
SKYKVLIIKQWLIKHLFNNQTDNILFSPHNACLSNYKLDENYVYLLIINNLANK
SEHLINIRKKNYTLDDHTNNYNSNTNIIITYSNNNNNNNNNNIILSDISYFNK
YLTKKIVENTSYLYICELLIIDIYLANTYGNLIYIKYKYDHENTYTNIIIS
IVQKKKEKKKSKYKNIYDHSKENNRKQSNISYISFSQKNYKNTNMNNNLHGE
EIPDSSSSNNISSSSSSSKSHNSATSYDIPKMRNRKNNLNYINRKNLNDLIY
FPKYKLIIPHTLNVNQNNSNIEAREILVLGIYRKIQKFIENKRSQCNFTED
IFCCLSNISHILHINRIEENVRNIYNTKINMKGNSTNNNNNNNNNNNNNNFI
HNDLSICLDFEYKLSCTDPEEKIIPFSINLPHKNTQDNFLNFINIEKAEISLY
KIKISVLPIQIGKLLKIIKKKIKNNKYLIDLIYNLQYLPNDIPDKKVLVY
LKNITPLFPFIKNIYIPYGVKDTYKFEFFIHNKHITPYKQNHKPYFNNWYKIFK
KYIDLSTNYMGAKYIILNDKVPKFKLSNHIITGKYIDVLSACSKILNTQKTSFY
LIDNEYKNINENIDNDNTQNNRNSFLPKDVIYKRIYENINNDNTSNDHRNP
YFYNDLYKQSTQGMNKQLINGKNTYNNLPEDEYIEHNNKDNKSKCFLIDYD
DLYEVLIHNNHLYASKOMPELYIKKIDIEKIKHYFFFLFOISDLYKYFVILSHE
DYEKLYHKNKNNIINKLKNFPDISRSVLHYLKYKDYVAISDIFSIMDNLIIID
LRKFOIQEQQKNGKANNYEVHVDGDDDDDDYNNNGNENYNNYEEGVNDEERNY
LNKNDYDEQADEEDDDDDDDDDDDDDDDDEYKEYENDNGNDYDEADSDHNNYNT
NENYHDDQVHNSNFIKNPLHMDNTNKKQKQHYNEEYKTKNERIIEKNAELNKG
IKKAKKLIKKKGNYSKLNELRRKNKNTNNNNNNNNNNNNNNNNNNNNNNNNNN
NNTFLKNNHNVSPNNVEYKGLIISYKNNFPYNNLFINNITIFKYTLFVRLNYCK
YIEHLSEVLMNHNFKVNIININECTNNLMICTHRECMIHFLKLYIHQNDVYKSE
YNNNTKLTFLIEDDIHINTYNNLILKYSFIMDNNIITCILKLSISHSIFRHL
KNFDKNEQIENITNHNKSVNLNPNTHIINHPHQYNNNNNNNNNNYHKKIIQER
LLRDKAYISLQNTIKHYDKFKNFPILLTEYVNNIIDEYAHNFKILKLDNYFTYNTK
KLISQNTSFNNDINSEYDNDFNDLDDSDNTQVTKENVNGVNNNNNNNNNNNNNN
SNIPIESNQVQIKRQKLLLMNDRYNNIDLNHNTNIDMNNINLMNKGNNYNYREN
ITSEIKHNSPNEYRNNIITNHTINNGNSYHQINYSLHRSNNNNNNNNNNNNKQYSL
HNLNDKSNFKKNYNTNDTNNNNHYVTNNNTSPKSSNEKNKIYNNHPINMYRNDYK
NILLNNDKDTILSENPLYKRNHIFPQNKHDTSPSLHINNFQDNNNNISPLNYSKUK
KFNITSYNNNNKNNINSLDLSIDNSRYSKNTA"
12790..13329
gene
CDS
/gene="PFC0166w"
/notes="synonym: fred"
12790..13329
/gene="PFC0166w"
/notes="Revised: split original spliced gene PFC0165w into
PFC0165w and PFC0166w, fasta; Plasmodium falciparum;
thioredoxin-like redox-active protein; fred; length 179
aa; id=100%; ungapped id=100%; E()=2.1e-69; 179 aa
overlap; query 1-179 aa; subject 1-179 aa; fasta;
TR:AP99466 (EMBL:AY003872); Plasmodium vivax;
pvlh14100_p; pvlh14100; length 179 aa; id=65.92%;
ungapped id=67.42%; E()=1.5e-46; 179 aa overlap; query
1-177 aa; subject 1-177 aa"
/codon_start=1
/product="thioredoxin-like redox-active protein, putative"
/protein_id="CAD49085.1"
/db_xref="GI:23477023"
/db_xref="GOA:Q81224"
/db_xref="InterPro:IPR006663"
/db_xref="UniProtKB/TrEMBL:Q81224"
/translation="MACQVNDPPKTPNDKTAIEYKAYVMNLYYYQNNELKKIDSS
YFKDKYLGLFGASMKCYCTFIDSLNIFKNFPNVEIYIPDRTYOEYOSFLKNTN
FYALPFDNYLYICKYQIKNLPFSMLITPNNNILLVDAALIKTDEYINNLSLIKNY
IITHPTFOFNRRPFDLERN"
complement (13855..15201)
/gene="MAL3P2.3"
/notes="synonym: PFC0170c"
complement (13855..15201)
/gene="MAL3P2.3"
/EC number="2.3.1.-"
/notes="Similarity to lipamide acyltransferases. eg
C.elegans lipamide acyltransferase (WP:ZK669.4) BLAST
Score: 582, sum P(1) = 9.2e-57; 33% identity in 430 aa
overlap; similarity to P. vivax pvlh14105C, 59.908%
identity (61.611% ungapped) in 434 aa overlap
Pfam match to entry PF00198 2-oxoacid_dh, 2-oxo acid
```

## Alignment Scores:

Pred. No.:	Length:	158548
Score:	Matches:	65
Percent Similarity:	Conservative:	42
Best Local Similarity:	Mismatches:	95
Query Match:	Indels:	73
DB:	Gaps:	14

US-10-734-719-9 (1-291) x PFMAL3P2 (1-158548)

QY	23	AsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyrTyrLeuGly	42
Db	66321	CATGATTATTATTATTTAGATAAATAATATATATTTTTTTT	66283
QY	43	LysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeu	62
Db	66282	-----AATGGTTTTTATACATATATATAATAAATA	66253
QY	63	LysHisLeuIleGlnAsnGlnGluTyrGluThrGluLeuIleMetCysSerAsnTyrAsn	82
Db	66252	AAA-----TTAAATAATACATCATATATAGAAGATATAAATTTATTTATCAATATTATG	66199
QY	83	GlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyrAspTyrPhePro-AspAl	102
Db	66198	ATAGGTTTTTTTCAATAAATAAATAATTCG-AAATTTATTTCTATGCTCTTTTATAATAATGT	66140
QY	102	ahisLeu-----GlyTyrAspPhePheLysGlnLeuLysGluPheAsnAlaTy	118
Db	66139	ACTACTACAGAAAAATACAGGGAATCATATATAAGGAATTA---AATTATAATTATGT	66083
QY	118	rPheLysPheHisGluIleTyrPheAsnGlnArgIleThrSerGlyValTyrMetCysAl	138
Db	66082	GTGTGCTTATATATAACCTTCTTTTCGAAG-----GT	66050
QY	138	aValAlaIleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyr-----	156

```

Db      66049  ATATGCTTATCAGAGGATACAAAAA---TATATGAAACATGTCAGCATTTTGTGAA 65993
Qy      157  ----GlnAsnGlySerTyrAlaPheAspThrLysGlnGluAsnLeuLeuLysLeuAl 175
Db      65992  AAGAAATAATTTAAATGATGATTGATTGTTCTAGAAAAGAAGATATGAAATAAGAGAGAA 65933
Qy      175  aProAspPheLysAsnAspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLy 195
Db      65932  AAAAGAAATAAAAAATAAAA-----GGTGATATGAATTTATACAAAT----- 65893
Qy      195  sAlaLeuGluPheLeuGlu--LysThrTyrLysIleLysLeuTyrCysLeuCysProAs 214
Db      65892  -AGTATAGAAATATCAGAGAGACAAACAAATAAA----- 65860
Qy      214  nSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsn----- 230
Db      65859  -GAGTATATAATCAATAAAATTAATAAATGATGAAATATATCCATAAGACAGATCTGAT 65801
Qy      231  -----PheIleIleGlnGlyLysAsnAsnTyr-----ThrLy 241
Db      65800  AATGTTGTATATTTATATATGCTCTAAAAATAATATTGTATATATATAATATTTTGAA 65741
Qy      241  sAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPheLy 261
Db      65740  ACATATTTTATTGATGTTATA-----AAATTTAAATAATTAATTAATTTTAA 65693
Qy      261  sLysIleLysIleLysGluAsnValTyrTyrLysLeuIle 274
Db      65692  CAAGATTAAATTAAGTAATTTCTTTTATAAGAAAAAT 65653

RESULT 140
AE014821
LOCUS   AE014821 250823 bp DNA linear INV 07-OCT-2002
DEFINITION Plasmodium falciparum 3D7 chromosome 14 section 6 of 13 of the
complete sequence.
VERSION  AE014821 AE014187
KEYWORDS
SOURCE   AE014821.1 GI:23497369
ORGANISM Plasmodium falciparum 3D7
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 250823)
AUTHORS  Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,M.J., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE     Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL   Nature 419 (6906), 498-511 (2002)
PUBMED    12368864
REFERENCE 2 (bases 1 to 250823)
AUTHORS  Gardner,M.J.
TITLE     Direct Submission
JOURNAL   Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
     source
       1. .250823
         /organism="Plasmodium falciparum 3D7"
         /mol_type="genomic DNA"
         /isolate="3D7"
         /db_xref="taxon:36329"
         /chromosome="14"
         147..196
         /rpt_type=tandem
       235..301
         /rpt_type=tandem
       386..449
         /rpt_type=tandem

repeat_region
repeat_region
repeat_region

```

```

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
gene
mRNA
CDS

LIGINTLICISCFKYPPEFVNKCVICNSFDMINLKEDKRNMNIVKKEIHIFMELY
KYKDFSHNSIFINLYNNMCKYLYYYETIHOQIVISNNFVLARTFPFSQLLOI
IOKDKHNMKNMREPHDDHEPDEEKYKOEIOIKTVTKENDVHNMHNSHL
SDHYNNIKONINHMVCVNTONEHLHDEVTNKVNEYOIKNFEGKNKKHIGBOTPEELN
SIQNDAQEENYIOEIKEDIMGERNNKKENYIMKWNTEKEKQIIOKEKMKHIEH
DTIIOKDDKVEKVENKVINVKDIDVKINVVNKDIDVEINVVNKDIDVKINMLD
NPNKDKINLNTVHNNKCKEDKKEECSEKRNKDILIVEKKNKDI FNKKKSINI
RNKLYKNKDKNSDSVENFFYTFSTKTIISHIEKLNKKIKFSFNFKNKTHIKLMELN
FNNTIRKYANGDILILYIKYLQREKEIHKLQNYKVIKKKKKKKKKK"
4753..4920
/rpt_type=tandem
5281..5336
/rpt_type=tandem
6760..6876
/rpt_type=tandem
7353..7386
/rpt_type=tandem
7397..7455
/rpt_type=tandem
<7677..>9677
/locus_tag="PF14_0303"
<7677..>9677
/locus_tag="PF14_0303"
7677..9677
/locus_tag="PF14_0303"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN36916.1"
/db_xref="GI:23497372"
/translation="MPKNMEVYKINBEFGLIEEVSQNNYTEKNNNINLVYKQNI
FRNDSLLRQKKENISINKLENVEYELLKENINSVLNILLFYDTFNKKNLNOLDLF
HILKKIMFLYDKELIKLYNHPNDQYNNHKNKTKSNEITKKKEKNVEKCDHLRS
NDIRPELICILINIKKNNNVLNNNKVQIINKLQGFNEPYNFDIUIYIIFKH
FISNKILNKYIHNFNKNVNVLEFNSIYNLKNKKKISFNSFHDYFNMVDTKNTIYI
KHQJLHKKEEKFNINNEENHSHINEYKINSVQHKCSQKCMWRKKKKKTSDDK
GNTLSLKNQKQSKYIKKKTLPFKKIINELNKLJOKYSEKYNILFNLYAOKKKK
IYNLKNCLFRNETLYQKNDIHKKEHKVEIMLVNNYDNTNHNKINEKIOEKDMKVC
KENEDSQDRKNLYDENNDIHTYDEDKTNPFLPSYGMKDINKINIKYINFNMMNK
FKKSIIFLQLEFLTSDENINIPKKKILHFTLVDINIPILYILKKNKNVYNFNEFK
KQNFNDIFINYOISYFNTNVNHNHIIHPLFKQNPYICFLLYIHDDFIPTNNEY
NKLHQHNKVENSRNMILQLLYIYMCNICECFNFIKQHC"
10309..10336
/rpt_type=tandem
complement(<10401..>15333)
/locus_tag="PF14_0304"
complement(join(<10401..12590,12766..>15333))
/locus_tag="PF14_0304"
complement(join(10401..12590,12766..15333))
/locus_tag="PF14_0304"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN36917.1"
/db_xref="GI:23497373"
/translation="WBEIEEKYIIIDALYDTWSDNDKRESEKILYEIEKDEKFI
CLFDIYTSKSIHYNIRKLGIIYCKNLIVRYNNRDGPHYSNTTKIKKKILDILNVV
EYLNRYEFSLLKRIGRYELVHNPBELLCLYNINIHKTNNINNIYIYLLYKILR
EQYSKLPKOKKETDISEKPIKSLFEFNNISLYNFQNNATDFCMCNDEKNNCK
NLNSILNLYKQLFVSNVDKTRNEGATANENIPEDENYIHDKKIHHMKYLDSDILNL
IINRNDIKVNNVYNSKMSKNGDQENKVDQYKVDQYKDAEYKVDQYKVDQYDAEY
KVQEPKVDQYKVDQYKVDSEYKECKENYKNKESPNSSNTTTNNNNNNNNNN
NVNVDNDLNTNFNLAFINKILLYLKFINNNSASYILYKLFKLFSTFLMIEFH
TSVHYLUKDIOTEKILCLFLKKNLSNNYNDYEHKEIIVNLCSILKISLYFHFCLN
QYNLIKQKRVENYINVKNNKNINENYNNVNTLESKIIINFMFINKSMENNNIL
FNKLSLKLDIYIREVQKALSNDPILIHNEKALFIYFLRYCINISSEQIIEVT
LNTRDYETRENTFYNNAKDCIIELSSEPFLFSIFHFEFPLINSFNHYVFLKGLPV
HLNKIQPESSAIIQDDGVLNIYIILYPSFHKIKIEHILCMIRFMDFLNIKYTHP
LISIRALIKNTWKTKYKTFPFEDDILLIENIRFLYIQLQKNVSVSLTENIPSY
VNDMTFOHYAKSNFNLPIILFFKPIFLSKYFFKYKYNDYINEFLVSLINVLTK
ITYPNKIQIKLISHLVFMSNKEKDIIPKDNYNFLDLDMNSNISIKYILNLIVQ.
ILANKRYEYGLSEYMKENYHNNKVGNEKVDVILFYPFSDIILYTLKAKIEK
OPKNDIIFNPHPIPSYKANHNINDITSIDNSVDNNINNFOPNNIYNLNPYI
KEINLDDKTSIDNFSIWLICILKLIYMKFHOKKEEIVKKICSYIATINLYEYFKN
QTGKISYELSNLCLDVVIEYMWCLFIHETHNHFYLTYESINIKMLTNDIISKTLKVV
```

```

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
gene
mRNA
CDS

EFNFILNDEIKIEKCIYIHLCLSIYKNDMKNEKSIYTYIAQFVYIYLLKVLKLYFS
KYFNFYIQQONKKAERNNDANKYNLHNSNNNNFNDVQKNGNVIFDYKYKTICEEK
SVNRKNIIPDISFDLQGNNTNNYLNHFNDNNNDNNNNNDNNNNNDNNNDNNNNNN
NNNNNNSSNSSEONNNKLNQCNKHIKVIYDNIINVOHDENFLNEFTHQNDVKKLFP
NINYINDYDITKTFLLNFENSTNEDNYSINKHSVFTLSILSYERYNYLIILITFL
TYPKINIPIFISFYQSQYIHNKNILNLSLIFIYILTQNYIILTFPFSIFLOQTFPSI
TNTQSFKMLYDNTNKNHFTCNKDEKTLIIIEILKLNINLNILNNDKIYIEKKNILHL
STISNLTGKTFHYSYNNVILRKVLKYDDFAS"
11177..11276
/rpt_type=tandem
12606..12658
/rpt_type=tandem
14155..14194
/rpt_type=tandem

Alignment Scores:
Pred. No.: 0.00543 Length: 250823
Score: 120.50 Matches: 70
Percent Similarity: 36.1% Conservative: 48
Best Local Similarity: 21.4% Mismatches: 110
Query Match: 7.7% Indels: 99
DB: 13 Gaps: 16

US-10-734-719-9 (1-291) x AE014821 (1-250823)

QY 17 AspTyrSerArgLeuProAsnAspPheAspValPheArg----- 29
Db 227354 GAATATTCTATGTTAATAAATTTTTCACATTTATATAAATAATATGTTTCAAAATTTTGT 227413

QY 30 -----CysAsnGlnPheTyrPheGluAspLys--- 38
Db 227414 TTTTACTACTACAATGGTATTACTGAATATATCAAGAAATATATGATGAGATAAAGGA 227473

QY 39 -----TyrTyLeuGlyLysLysCysLysAlaValPheTyrThr----- 51
Db 227474 GAATATATTATATGATAAAGAAATAATATATGATATATACATATGAGATAAATAATG 227533

QY 52 -----ProAsnPhePhePheGluGlnTyrTyrThrLeuLysHisLeuLeGlnAsn 68
Db 227534 AAACATATTATTAATATAAATAGACGATATATTATAAATCAACATCGATTATATGAT 227593

QY 69 -----GlnGluTyrGluThrGluLeuLeu----- 76
Db 227594 ATAAATAAATTTACTATATATTATAAATATATGAAAAATATCTTTTCAATTTTATTACT 227653

QY 77 -----MetCysSerAsn-----TyrAsnGlnAlaHisLeuGlu 87
Db 227654 TATAACATATATAAATGTTTCCAATATACAAATGAAAGAGATAAATGAAAAATAAAGGAA 227713

QY 88 AsnGluAsnPheValLysThrPheTyrAspTyrPheProAspAlaHisLeuGlyTyrAsp 107
Db 227714 AATGAGATTTTAAAGAAATTCCTTTTATTATATATATATATATATATATATATATATT 227773

QY 108 PhePheLysGlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsn 127
Db 227774 ATTTTATAAAAAAATAAATAATAT-----CATATAATTCCTTTGAAAC 227818

QY 128 GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGlu 147
Db 227819 AAT-----GTAGAATTAATTTTAAATATATAGTAAG 227848

QY 148 IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerTyrAlaPheAspThrLys 167
Db 227849 CAAATTATAAAAAAGATAATATATATTTTAAAGTCTTAATTTATGGGCATCTC----- 227899

QY 168 GlnGluAsnLeuLeuLysLeuAlaProAspPheLys---AsnAspArgSerHis----- 184
Db 227900 -----GGATATATACAAATAATGATGACTTTAAATGATGATGATGATGATGATGATGAT 227953

QY 185 -----TyrIleGlyHisSerLysAsnThrAspIleLys----- 195
Db 227954 ATTAATAATTTTCAAAATGCTTAAACATACATAAATAAATAATTAACAAATGAAGTGAATGT 228013
```

Qy	196	-----AlaLeuGluPheLeuGluLysThr-----TyrLysIleLeuLysLeuTyrCysLeu	211	rpt_type=tandem
Db	228014	AATGTATTATCAACACCATTTTATGAAAGACAAATTTGATATATATTATTATTTGAAAAA	228073	repeat_region
Qy	212	CysProAsnSerLeuLeuAlaAsnPheIleGluLeuAlaProAsnLeuAsnSerAsnPhe	231	repeat_region
Db	228074	TTAAGAGATAAGAAATTTGCACCTATTATAAATATTATTTCAACATTTATATAAATTC---TAT	228130	repeat_region
Qy	232	IleIleGluGluLysAsnAsnTyrThrLysAspIleLeuLeuProSerSerGluAlaTyr	251	repeat_region
Db	228131	AATAATAAGATATACGAATTTACAAATGACATAATA-----	228169	repeat_region
Qy	252	GlyLysPheSerLysAsnIleAsnPheLysLysIleLysIleLysGluAsnVal-----	269	repeat_region
Db	228170	-----CCTTATAATATTACCAATTCATATTAAATTTTGAATATAATAGTGTTCCT	228220	repeat_region
Qy	270	---TyrTyrLysLeuIleLys	275	repeat_region
Db	228221	TTATTTTATCATATCTTCAA	228241	repeat_region
RESULT 141				
AE014844/c				
LOCUS	AE014844	251551 bp	DNA linear INV 11-FEB-2003	
DEFINITION	Plasmodium falciparum 3D7 chromosome 12, section 1 of 9 of the complete sequence.			
ACCESSION	AE014844	AE0141188		
VERSION	AE014844.1	GI:23496531		
KEYWORDS	Plasmodium falciparum 3D7			
SOURCE	Plasmodium falciparum 3D7			
ORGANISM	Plasmodium falciparum 3D7			
REFERENCE	1 (bases 1 to 251551)			
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. Gardner M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallow, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B.			
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum			
JOURNAL	Nature 419 (6906), 498-511 (2002)			
PUBMED	12368864			
REFERENCE	2 (bases 1 to 251551)			
AUTHORS	Hyman, R.W., Fung, E., Conway, A., Kurdi, O., Mao, J., Miranda, M., Nakao, B., Rowley, D., Tamaki, T., Wang, F. and Davis, R.W.			
TITLE	Direct Submission			
JOURNAL	Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
REFERENCE	3 (bases 1 to 251551)			
AUTHORS	Hyman, R.W., Fung, E., Conway, A., Kurdi, O., Mao, J., Miranda, M., Nakao, B., Rowley, D., Tamaki, T., Wang, F. and Davis, R.W.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
FEATURES	Location/Qualifiers			
source	1. .251551			
	/organism="Plasmodium falciparum 3D7"			
	/mol_type="genomic DNA"			
	/isolate="3D7"			
	/db_xref="taxon:36329"			
	/chromosome="12"			
	4. .647			
repeat_region	/rpt_type=tandem			
repeat_region	656. .797			
repeat_region	/rpt_type=tandem			
repeat_region	754. .933			
repeat_region	/rpt_type=tandem			
repeat_region	775. .807			
repeat_region	/rpt_type=tandem			
repeat_region	2009. .2796			





CDS	ROEYNKGDIIILFNLNINNIYGLL" 9950..10315 /note="RPL14" /codon_start=1 /transl_table=11 /product="ribosomal protein L14" /protein_id="AAO40229.1" /db_xref="GI:31322463" /translation="MTQINTVFNIDNSGVKKILCIQNLTKTKKIEIGDLIVGVKK INNTSKLYSNIYGIIVRLKKNINLYKKNISFNDNSAVLVDRNLNPIGSRIFGTIP KFLKKNYKLYSLTVDFI" 10317..10676 /note="RPS8" /codon_start=1 /transl_except=(pos:10464..10466,aa:Trp) /transl_table=11 /product="ribosomal protein S8" /protein_id="AAO40230.1" /db_xref="GI:31322464" /translation="MILKKNILTNININKYKKYLIIISYTKITYLFFILLKNFYIKNI YKLENILVYIYSKKLKPKFFNFKDKKNITYLQKKIKKQNKNGIIFTSKGIFFI SEAIKQGGILFLDIY" 10701..11249 /note="RPL6" /codon_start=1 /transl_table=11 /product="ribosomal protein L6" /protein_id="AAO40231.1" /db_xref="GI:31322465" /translation="MKQIVKEQLINSFNLFFSILGIKTQIIVLFILYNEKYQVIFLSN KIISPPGFYLYLPSKKNIINTYINTLKNKILFCIKYKLSIEPIGIGYKFLKKNF LYFLGYSHLIKILPNKNCIYIDNNKILNLVSNKYILGNIFYIISKCNLNVYKKG GIRNIFQLKLGKPKRNKNK" 11249..11944 /note="RPS5" /codon_start=1 /transl_except=(pos:11684..11686,aa:Trp) /transl_table=11 /product="ribosomal protein S5" /protein_id="AAO40232.1" /db_xref="GI:31322466" /translation="MIFLNKKIYKNKYIYKKNINLLYFIKIINDPIFSPFLIKYIKI YNLSPFNLLKNFINLLAIKYLNIYNNIPNITTIYNIKIYLEIFYWENKNFIN YNYKLLFTTEQKILLIKVSKTRAKRIKKYKIIIIIGNKSGWFGICSDYILODA ISKRLHAFKNIYQISLFYSNLLKNNIYIKKKPKLYLFSYNYKIKISSYIIRLLFY FIGINNFSKIMG" 11991..12140 /note="RPL36" /codon_start=1 /transl_table=11 /product="ribosomal protein L36" /protein_id="AAO40233.1" /db_xref="GI:31322467" /translation="MKVKLSIKRLCKNGFLINRQKKYIYGVKFLNITTDKNNKNLNI VILII" 12145..12543 /note="RPS11"	Db	20984	TTAATTAATAATATATATACCTAAATATAT-----TAAATCGATAATTTATTT	20937
		Qy	71	TyrGluThrGluLeuile	76
		Db	20936	TATAAACCAATTTTAAATACAAATAAATTTTATAGATATAAATTTATTTTATAAAAAAT	20877
		Qy	77	-----MetCysSerAsn	80
		Db	20876	ATAATAATTCCTAAAAATATAGTATCGCAAAATATTACACAGACAATTTATAAATAATAT	20817
		Qy	81	TyrAsnGlnAlaHisLeuGluAsnGluAsn	90
		Db	20816	TTTAAATATATTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	20757
		Qy	91	-----PheValLysThrPheTyrAspTyr	99
		Db	20756	TATATATATAAAAAATATATATATATAACACCAATATTATATCTTATTTGAAAAATTTATTT	20697
		Qy	100	ProAspAlaHisLeuGlyTyrAspPhePheLys	110
		Db	20696	TCAATAAATAATATA-----TTTAAATTTTTTTAGAGGAATTTTAAATAAAAAAATAT	20640
		Qy	111	-----GlnLeuLysGluPheAsnAlaTyrPheLysPheHisGluIleTyrPhe	126
		Db	20639	TATTTCTAACCAATTCATTAAAGATATTACATCAGGATTAATTTCTTATAGAAATTTATTTT	20580
		Qy	127	AsnGlnArgIleThrSerGlyValTyrMet	140
		Db	20579	GAAGCTAAAAATATACCTAATATATATCTTTTACCTCAAAAATAAATATTTTAAACAAAT	20520
		Qy	141	IleAlaLeuGlyTyrLysGluIleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySer	160
		Db	20519	ATATTTTATTTATATGAAGAA-----AATACATTAAATTTATTTATGAAATTTATAGT	20469
		Qy	161	SerTyrAlaPheAspThrLysGlnGluLeuLysLeuAlaProAspPheLysAsn	180
		Db	20468	TTATATACTTTTCTTATAATAAAAAAATTTATTTTTCACAGAA-----ATAAAT	20415
		Qy	181	AspArgSerHisTyrIleGlyHisSerLysAsnThrAspIleLysAlaLeuGluPheLeu	200
		Db	20414	TCATTATCAAATTTATTT-----	20397
		Qy	201	GluLysThrTyrLysIleLysLeuTyrCysLeuCysProAsnSerLeuLeuAlaAsnPhe	220
		Db	20396	-----TATGAAAAATAAACATCTATATTAAACACCTAAT-----	20364
		Qy	221	IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr	240
		Db	20363	-----AATATTTATTT-----CCTGGTAAATTTATTTCT	20337
		Qy	241	LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe	260
		Db	20336	ATAAATATTTTATTA-----TTAAATTTATTTTTCACATAATTTAAATTTT	20292
		Qy	261	-----LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAsp	276

Alignment Scores:

Pred. No.: 0.000933 Length: 34750  
Score: 120.00 Matches: 70  
Percent Similarity: 33.9% Conservative: 43  
Best Local Similarity: 21.0% Mismatches: 92  
Query Match: 7.7% Indels: 128  
DB: 13 Gaps: 15

US-10-734-719-9 (1-291) x AY217738 (1-34750)

Qy	31	AsnGlnPheTyrPheGluAspLysTyrTyrLeuGlyLysCysLysAlaValPheTyr	50
Db	21044	AATAATTTAATAAGATCTTATAGTATTATAAAAAAATAAATCTTTTAT	20985
Qy	51	ThrProAsnPhePheGluGlnTyrTyrThrLeuLysHisLeuIleGlnAsnGlnGlu	70



REFERENCE	Orthopoxvirus.		
AUTHORS	1. (bases 1 to 189274) Jin,Q., Hou,Y.D., Cheng,N.H., Yao,E.M., Cheng,S.X., Yang,X.K., Jing,D.Y., Yu,W.H., Yuan,J.S. and Ma,X.J.	CDS	NHLGIGVKDKRISDSFGDVCMDDTIFDPSEARELPSLTNDNRRNIAMDTKLDDDDTDI WTPVTENDYKFLSLRLVYAKSQSDIVFDYVVLVTGDEPTVTFIKVTRFYFNNPK" 8294..8491 /note="ORFR" /codon_start=1 /product="unknown" /protein_id="AAF33859.1" /db_xref="GI:6969647" /translation="MSTHSTFFNRLYSIFLIFSISLSESSICILSESVTCLNLSLM SFLQSNLSLANSILVCRPLSS" 9014..9148 /codon_start=1 /product="TC16R" /protein_id="AAF33860.1" /db_xref="GI:6969648" /translation="MEFDPAKINTSSIDHVTILOYIDEPNDIRLTVCIQNIISISOK "
TITLE	Complete genomic sequence of vaccinia virus (Tian Tan strain)		
REFERENCE	Unpublished		
AUTHORS	2. (bases 1 to 189274) Jin,Q., Hou,Y.D., Cheng,N.H., Yao,E.M., Cheng,S.X., Yang,X.K., Jing,D.Y., Yu,W.H., Yuan,J.S. and Ma,X.J.	CDS	complement (9843..10049) /codon_start=1 /product="TC15L" /protein_id="AAF33861.1" /db_xref="GI:6969649" /translation="MKSNEDETFIEHLGDGIKEDETVRTTDSGITTLRKVLHVTDNPK FAHVRFTCLVTTIDGVSKNIWLK" complement (10567..10995) /codon_start=1 /product="TC14L" /protein_id="AAF33862.1" /db_xref="GI:6969650" /translation="MSYACPILSTINICLPYLKINDIMDKRGETLLHKA VRYNKQSIV SLLLESQSDVNIIRNSNGYTCIAIAINESRIELLMKLCHKPTLDVYIDSLREISNIV DNYAIOKICIKYAMIIDCTSSKIPESISQRIYNDYIDLN" complement (11084..11431) /codon_start=1 /product="TC13L" /protein_id="AAF33863.1" /db_xref="GI:6969651" /translation="MADYINSDYRNKVDLDLVLFPQNGKPHGIMCSIVPLWRNDK KPSFLILKTMNSDVLQHLIEYMTFGDIPLVEYGTVVNKEAHEYFRNINIDSYTMKY LLKKEGRCHQLSR" complement (11762..11995) /codon_start=1 /product="TC12L" /protein_id="AAF33864.1" /db_xref="GI:6969652" /translation="MLKLDIAMALLEATGFNSINDFNIFS YMKSKNVDVDLIKVLVE HGFDLSVKCENHRSVIENYVMTMILFLKLLICS" complement (12001..12231) /codon_start=1 /product="TC11L" /protein_id="AAF33865.1" /db_xref="GI:6969653" /translation="MFDYLENEVALDELKOMKORDPNDTRNQFKNALHAYLFNEH CNNVEVKLLDSDGTNPLRKIGDSYPIRIHK" complement (12358..12537) /codon_start=1 /product="TC10L" /protein_id="AAF33866.1" /db_xref="GI:6969654" /translation="MSSKGGSGGMSVFIHGHGSGNKGTKTYTSGGGGMMGGGSSGV KSGVNGGVKSGTGKI" complement (12710..14614) /codon_start=1 /product="TC9L" /protein_id="AAF33867.1" /db_xref="GI:6969655" /translation="MVNDKILYDSCKTFNIDASSAQSLIESGANPLYEDGCTPLKAY VTKNKNIKNVDVILLSSVDYKINDFDIFEYLCSDNIDINLLKLLISKGIEINSIK NGINIVEKYATTSPNPDVFKLLDKGIPTCNSNIQYKIKIEQIRAGEYNNVDEL DDYDYVTTDDRMKTVLYYIITRSODGYATSLDINLYLISHEKMRYYTYRHT TLYYLDKCDIKREJFDALFDSNSYSGHELMNLSNVLKQFKKNHKIDNYIWDQLLF DRDTFYILELNSLRNNLISILKRYTDSIQDLELLEYSYHTVINVLKCMDEGAT LYRFKHINKYFQKGNRPKVVEYILKGNLVVDNDNDNLDNINIMLPFTFMSKRELDV
TITLE	Submitted (29-SEP-1998) The State Key Laboratory for Molecular Virology and Genetic Engineering, Institute of Virology, 100 Ying Xin Jie, Xuan Wu Qu, Beijing 100052, P.R. China		
JOURNAL	Location/Qualifiers		
FEATURES	1..189274 /organism="Vaccinia virus Tian Tan" /mol_type="genomic DNA" /strain="Tian Tan" /db_xref="taxon:10253" complement (4274..5008) /codon_start=1 /product="TC20L" /protein_id="AAF33853.1" /db_xref="GI:6969641" /translation="MHVPASLQSSSSSSCTEBENKHMIGDILVILKQDOTPTND KICQSVETIESEDPDEVEDSTSVEDPPTTYYSIIIGGLRNFQFTKCPQI KSISESADGNTVNARLSPGQKSDIPATHEALAMIKDCVYSIDIRCSEEEKSD IKTHPVLGNSI SHKKVSYEDIIGSTIVDTKVRNLEFSVRIGDMCKSELEVRDGEK YVDSASEGATDDTSLIDSTKLKACV" 4490..4693 /note="ORFR" /codon_start=1 /product="unknown" /protein_id="AAF33855.1" /db_xref="GI:6969643" /translation="MRDVRBYWGLDVAVLFFATSDVDRHLTVFDHKSFFWMSDRG RVITLSWGHAGQSSIHCVSISGP" 4497..4610 /note="ORFR" /codon_start=1 /product="unknown" /protein_id="AAF33856.1" /db_xref="GI:6969644" /translation="MLDPSTGWLMSLSFSSSLHMSIDTQSLLIIRASS" complement (5444..6505) /codon_start=1 /product="TC19L" /protein_id="AAF33854.1" /db_xref="GI:6969642" /translation="MDIFKELIVKHPDENVLISVLSILSTLSILNHGAAGSTABQLSK YIENMNTPDDNMDVDIPYCATLATANKIYGSDSIEFHASFLOKIDFQTVNFN NANOTKELINWMTNGKINSILTSPLSINTMTVSVAVHFAMWKYPFSKHLTYT DKFVSKNI VTSVDMVGTENNLOYHINELFGGFSIIDIPYEGNSMWIILPDDIEG IYNEKNITBEKFKWCGMLSTKSIDLYMEKFKVEMTEPNLPILLENLGIINTLPGYY ADFSKMCNETIVTEKFLHTPIFDVNEBYTEASAVTGVMTNFSWYRTKVITNHPFYI MIKDNGRIILFIGRYCYPQ" 6927..7349 /codon_start=1 /product="TC18R" /protein_id="AAF33857.1" /db_xref="GI:6969645" /translation="MLINYLMLLFAAMIIRSFADSGNAIETTLPEITNATTDIPAIRL CGEGDGYCLHDIDIRGCRSHGYTGIRCOHVLVDYQRSEKPNNTTTSYIP SPGIMLVLGIIITCCLLSVYRFRTRTKPLQDMVVP" complement (7504..8499) /codon_start=1 /product="TC17L" /protein_id="AAF33858.1" /db_xref="GI:6969646" /translation="MDIYDDKGLQTIKLENNPEPCIRNDIRELFKHVTDSDSIOLPME DNSDIENIRKILYRLKNVCEVDISTIFPMKYDPDNDKRTCSNWVPLTNNWMEYC LVYLETPICGGKIKLVHPTGNIKSDKDIIMFAKTLDFSKTVLGRKTIPLAVLDSVSY NRSMTTHYNDVDIDIDHTDKNGKELCYCIVITIDHVLVDVTETIGVIVNRSKGCLLVN	CDS	

```
LSILKCPYIDDIINKIDHGCSILYHCICKSHSVSLVLEWLIDNGADINIITYGFTCI
KICVLADKYPEIAELVYKILEILSLKPTIECKTKTVYDLDHRVLFPGNNKSL
KICVFLVDYKYTCMSVPSVIEFTDCEKEIADMROJIKINGTOMLVYVGNKPKTK
KRYVNNPIFTDWANKQYFYNGIINANKLIBEQSKIDDMIEVSDIDNRLSTLPLEI
RHLPSYAPL"
14208..14420
/note="ORFR"
/codon_start=1
/product="unknown"
/protein_id="AAF33868.1"
/db_xref="GI:6969656"
/translation="MLLVHGIPLSNNSLNTSTLGFVDVWAYFSTILPILFISIPFE
ISNFKFISMLSEHRSYNSISLIPL"
complement(14657..15190)
/codon_start=1
/product="TC8L"
/protein_id="AAF33869.1"
/db_xref="GI:6969657"
/translation="MSSIRFTIACLGILFVLGNCHDDPPYQPDFKNIITLDIVTYEDLV
PYVDNDTTSFKVIYFKKFWITVMTKCAPFDITVSVVTSHDNLNIQFYSRDEYDOS
EDKICTIDVKARCNHLTKPEVTVOKEARYISLSSDLSCFDSIDLIDLIETNSTDTTV
LKSVELMLPKRAKSIHN"
14667..14915
/note="ORFR"
/codon_start=1
/product="unknown"
/protein_id="AAF33870.1"
/db_xref="GI:6969658"
/translation="MDLARLGSMSSYDFSTVSVLLVSIIRSMRSRSTESKHDRSEDNEY
LVASPTVTSGFVRWLHRLTSMVOILSSLCSVSYSL"
14840..15064
/note="ORFR"
/codon_start=1
/product="unknown"
/protein_id="AAF33871.1"
/db_xref="GI:6969659"
/translation="MVASCNVGNFNLALCIIIFVPTIKLYIQIIMRCVYANGINKR
STPFSHNSNPFPFKYLNERSCVIYSVWY"
complement(15262..15714)
/codon_start=1
/product="TC7L"
/protein_id="AAF33872.1"
/db_xref="GI:6969660"
/translation="MGIOHEFDIINGDIALNLQHKGDNYGCKLIISNDYKCLKFP
RFIIRPDWSEIDENVGLVFANNYAVKNKVDFTFYIYEAVIHLNKKTEILLYSD
DENELFKHYPIYISLNMISKYKVEENYSPIEHLIPYRDYESMD"

Alignment Scores:
Pred. No.:      0.0048      Length:      189274
Score:          120.00      Matches:      62
Percent Similarity: 38.4%      Conservative: 36
Best Local Similarity: 24.3%      Mismatches: 83
Query Match:      7.7%      Indels:      74
DB:              10      Gaps:      16

US-10-734-719-9 (1-291) x AF095689 (1-189274)

QY 4 ValilleleAlaGlyAsnGlyProSerLeuLysGluIleAspTyrSerArgLeuProAsn 23
Db 179650 GTACTGATTGAATGGGAAATCCATCCGGATCGCTTATAGGATTC----- 179694

QY 24 AspPheAspValPheArgCys-----AsnGlnPheTyr 34
Db 179695 GATTTTGATGTACTCTGTTTAACTAGTAGAGCGGTATCACCGAGCGACCTTGATC 179754

QY 35 PheGluAsp-----LysTyrTyrLeuGlyLys-----LysCysLysAlaValPheTyr 50
Db 179755 TTTGAAATGTACTGAAGAATATATAGGTAATACATATAAATGTCGTGA----- 179805

QY 51 ThrProAsnPhePhePheGluGln-----Tyr 59
Db 179806 ---CACAACTATATTTTGAAAAAACCCTTACAACTACAGTAGTATTGGAGTAATATAC 179862
```

```
QY 60 TyrThrLeuLysHisLeuIleGlnAsnGlnGluTyr-----GluThrGluLeuIle 76
Db 179863 AATGCATTTTATATACATTACTGAAATATTATTACTGAATATTATTACTGAATATTATTA 179922

QY 77 MetCysSerAsnTyrAsnGlnAlaHisLeuGluAsnGluAsnPheValLysThrPheTyr 96
Db 179923 TTAATTATATCGTATTGTTGCTGTATAGAAATGAAGATACGCGACTTACTAGG---TAT 179979

QY 97 AspTyrPheProAsp---AlaHisLeuGlyTyrAspPhePheLysGlnLeu----- 112
Db 179980 TTGTATCTCCCGATAGAGAACATATAAATGTAGACTCTATTAAACAGTTGTGTAATAA 180039

QY 113 LysGluPheAsnAlaTyrPheLys-----PheHisGluIleTyrPheAsn 127
Db 180040 TCAGATCCTAATGCATGTTATAGATGTGGATGTACGGCTTTACATGACTCTTTTATAAT 180099

QY 128 GlnArgIleThrSerGlyValTyrMetCysAlaValAlaIleAlaLeuGlyTyrLysGlu 147
Db 180100 TATAGATCAGTCAACGCGAAATACAGTATAGATACAAC-----GGTTACTATCAA 180150

QY 148 IleTyrLeuSerGlyIleAspPheTyrGlnAsnGlySerSerTyrAlaPheAspThrLys 167
Db 180151 TATTATTATCT-----ACCGATTATGAAATTAATGAATATTATTATGATGATTAT 180204

QY 168 GlnGluAsnLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGly 187
Db 180205 GATAGAACTGGTATGAACAGTGAGATGAGCGAGAGTGAT----- 180246

QY 188 HisSerLysAsnThrAspIleLysAla-----LeuGluPheLeuGluLys 202
Db 180247 -----AATATATCAATCAAAACAGATATGAGAATGAATGAATCTTCTATGATGAA 180297

QY 203 Thr-----TyrLysIleLysLeu 208
Db 180298 ACACAAGATCAAAAGTACACAACTAGTAGTTACGACATTAAATC 180342

RESULT 144
AE014818/c
LOCUS AE014818 252632 bp DNA linear INV 07-OCT-2002
DEFINITION Plasmodium falciparum 3D7 chromosome 14 section 3 of 13 of the
complete sequence.
ACCESSION AE014818 AE014187
VERSION AE014818.1 GI:23497188
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 252632)
AUTHORS Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
Kyes,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J.,
Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,
Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
Barrell,B.
TITLE Genome sequence of the human malaria parasite Plasmodium falciparum
JOURNAL Nature 419 (6906), 498-511 (2002)
PUBMED 12368864
REFERENCE 2 (bases 1 to 252632)
AUTHORS Gardner,M.J.
DIRECT SUBMISSION
TITLE Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
location/Qualifiers
1..252632
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="14"
```



```

source
1. .27273
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="14"
75. .96
repeat_region
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(113. .148)
repeat_region
/note="AT rich"
/rpt_type=tandem
complement(157. .235)
repeat_region
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
196. .235
repeat_region
/rpt_type=tandem

```

```
repeat_region complement(244. .308)
/notes="AT rich"
/rpt_type=tandem
317. .353
/rpt_type=tandem
/rpt_type=tandem complement(317. .353)
/rpt_type=tandem /rpt_unit_seq="(TA)n"
complement(400. .532)
/notes="AT rich"
/rpt_type=tandem
/rpt_type=tandem complement(534. .559)
/rpt_type=tandem /rpt_unit_seq="(TA)n"
799. .822
/rpt_type=tandem
/rpt_type=tandem complement(823. .947)
/notes="AT rich"
/rpt_type=tandem
1009. .1060
/rpt_type=tandem
/rpt_type=tandem complement(1012. .1099)
/rpt_type=tandem /rpt_unit_seq="(TA)n"
1012. .1059
/rpt_type=tandem
complement(1182. .1427)
/notes="AT rich"
/rpt_type=tandem
complement(1428. .1571)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
1580. .1607
/rpt_type=tandem
1580. .1606
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(1607. .1627)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
complement(1852. .1964)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
1856. .1948
/rpt_type=tandem
complement(2052. .2092)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
complement(2216. .2245)
/notes="AT rich"
/rpt_type=tandem
complement(2288. .2338)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
2341. .2500
/rpt_type=tandem
/rpt_unit_seq="(TAAAA)n"
2517. .2543
/rpt_type=tandem
complement(2517. .2543)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(2554. .2585)
/notes="AT rich"
/rpt_type=tandem
complement(2591. .2614)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(2625. .2656)
/notes="AT rich"
/rpt_type=tandem
2662. .2690
/rpt_type=tandem
complement(2662. .2690)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(2778. .2807)
/notes="AT rich"
/rpt_type=tandem
complement(2808. .2856)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
2808. .2852
/rpt_type=tandem
complement(2870. .2929)
/rpt_type=tandem
/rpt_unit_seq="(TAAAA)n"
complement(2966. .2987)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(3049. .3070)
/notes="AT rich"
/rpt_type=tandem
3078. .3150
/rpt_type=tandem
complement(3078. .3142)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
complement(3206. .3454)
/notes="AT rich"
/rpt_type=tandem
3297. .3376
/rpt_type=tandem
complement(3297. .3376)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
complement(3481. .3504)
/notes="AT rich"
/rpt_type=tandem
complement(3570. .3597)
/rpt_type=tandem
/rpt_unit_seq="(A)n"
complement(3654. .3692)
/notes="AT rich"
/rpt_type=tandem
3811. .3859
/rpt_type=tandem
/rpt_type=tandem
/rpt_unit_seq="(GAAAA)n"
complement(3868. .3928)
/notes="AT rich"
/rpt_type=tandem
complement(3949. .3980)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
3949. .3977
/rpt_type=tandem
complement(4011. .4056)
/notes="AT rich"
/rpt_type=tandem
4187. .4288
/rpt_type=tandem
complement(4187. .4288)
/rpt_type=tandem
/rpt_unit_seq="(TA)n"
4187. .4282
/rpt_type=tandem
complement(4290. .4339)
/notes="AT rich"
/rpt_type=tandem
complement(4356. .4391)
/notes="AT rich"
/rpt_type=tandem
4395. .4436
/rpt_type=tandem
complement(4395. .4436)
```

JOURNAL

The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus,  
Hinxton, Cambridge CB10 1SA, UK  
For more information about this sequence or the Malaria Project,  
see [http://www.sanger.ac.uk/projects/P\\_faiciparum](http://www.sanger.ac.uk/projects/P_faiciparum).

## FEATURES

## source

1. 341050

/organism="Plasmodium falciparum 3D7"

/mol\_type="genomic DNA"

/isolate="3D7"

/db\_xref="taxon:36329"

/chromosome="7g"

/gene="PF10780w"

969..1469

/gene="PF10780w"

969..1469

/codon\_start=1

/product="hypothetical protein"

/protein\_id="CAD51842.1"

/db\_xref="GI:23505060"

/db\_xref="UniProtKB/TREMBL:Q81223"

/translation="MELNKNEKKKKDDDEWELHPLFLSKIPKQNDIKNAALSALIT  
LINEEKEEYIEPRKRLKRIKTAGETIHKORRMYPQYQNDIKCFCKNDP  
NKDNLINDNTQIILNVNHNIDQNIYIENNNTSNPDTTNTNRKDGSEISMGELLVCL  
SMNLK"

complement(join(2246..2305,2416..2485,2589..2748,

2867..6716))

/gene="PF10785c"

complement(join(2246..2305,2416..2485,2589..2748,

2867..6716))

/gene="PF10785c"

/note="asparagine-rich C terminal; Pfam match to PF00083,

sugar (and other) transporter

11 probable transmembrane helices predicted for PF10785c

by TMHMM2.0 at aa 848-870, 877-896, 901-923, 936-958,

978-1000, 1163-1185, 1200-1222, 1243-1265, 1269-1291,

1304-1326 and 1341-1360"

/codon\_start=1

/product="transporter, putative"

/protein\_id="CAD51843.1"

/db\_xref="GI:23505061"

/db\_xref="GOA:Q81222"

/db\_xref="InterPro:IPR007114"

/db\_xref="InterPro:IPR011701"

/db\_xref="UniProtKB/TREMBL:Q81222"

/translation="MAEIKNLIKKEKNIYKSKIKNNKNSNNIYINKNINIYNNKNI  
HIYNEHNKYKEPEIKQNVKYIHSADTPREVYPKNYIISIINSEKRNKMEKXKKA  
GSOEIOTGIEYNNKNDNNNDNGDKINGDINADNNDNIYNNIHRQKSSNN  
SNKSSHLVHSNEMKNNKQIHWNVCTNNSEYNTNNIIEKCKDKCYLLYNNK  
DSYVYNGQNNINIYNNYNNIIEKGNHNNIYNTKPVDEYHYSKLOPVIINNKN  
NNNNNNNNNNNNNDNCINIISKPNYKQNLKNNISVVDQKKKTNNNIICS  
SDFVKSNNIYPTIQVNEKKYIASKYKSPSIKNNNVYLKKNKSIYINNKTPEYST  
NCVIKANNIKYDHNDFNNIYNNYHNSNNNNINENTTLCTNNKIYEHFYSNN  
AKTELKRETRHNDKQVONCKLYSDYNDKYVEDIQNNLKQKKNINIYENMLNKKR  
EGNNSTENFNNNNSKYKKEEDNTVKNIDKVTYEPNNYDNVILPKKNESTN  
FYNNHNNINI KKKNLKSSYNNKNNIILIQPEDKQVLPSELINSVEDMSDYI  
ESTPSHNSUSKIPENIMFQNSNTICHSKSGKYSDSYNDENTNNINVEIMDSYN  
FRKNNIYNNKKEEDIKNNIKLSDFMHDVNDKDNENKSKYQKASISYGLLTS  
FNKSKSEHILTLTSLUMI C I S I L C N D H G A I P V T L E E I O K D F P L S Y I E O S L G S L V Y  
MNNKILTQONERIKESNNCINISWGIDKYDTLLQIETDENTENNNDGNVITLKLLEY  
VLLKTLKLVKLLFHNKLYIITILGMSNLYFVVTQIOWITEYMSVILLTEKMKIVTVS  
TLCLFTSPTSGWFGFVCDLFGCYKNTNYSKTIKATATAFASACIIFGILSHLKNFI  
FFSISLWCLTSGALPFCVGMULSCVNNHQSLSASVQVITNVFQWFSAPLLSGI  
IMDITMHYTDNRALKAGETMIIYSSISIGFLFLFYANFLDLSKKNDDDILEEPL  
T"

join(9491..9535,9582..9722,9830..9973,10080..10148)

/gene="PF10790w"

join(9491..9535,9582..9722,9830..9973,10080..10148)

/gene="PF10790w"

/note="ProfileScan hit to PF50223, Thioredoxin-domain  
(does not find all). Similarity to Plasmodium falciparum  
thioredoxin trX TR:Q9NFK9 (EMBL:AJ277839) (104 aa) fasta  
scores: E(): 9.3e-06, 29.670% id in 91 aa. Similar to

Gallus gallus thioredoxin txN SWALL:THIO.CHICK

id in 105 aa, and to Ophiophagus hannah thioredoxin txN

SWALL:Q98TX1 (EMBL:AF321769) (105 aa) fasta scores: E():

1.3e-08, 36.44% id in 107 aa"

/codon\_start=1

/product="thioredoxin, putative"

/protein\_id="CAD51844.1"

/db\_xref="GI:23505062"

/db\_xref="GOA:Q81221"

/db\_xref="InterPro:IPR006662"

/db\_xref="InterPro:IPR006663"

/db\_xref="UniProtKB/TREMBL:Q81221"

/translation="MKCYEMINIHINIYEQSIYIELKNTGSLNQVSSSTQNSSIV

KFGAVWCKPCNKIKKEYFKNLNYYVTLDVIDIHPLKNDQHNIKALPTFPFYNLN

NEWLVHTVEGANQNDIEKAFQCYCLEKAK"

join(10913..11086,11363..14623)

/gene="PF10795w"

join(10913..11086,11363..14623)

/gene="PF10795w"

/note="Signal anchor predicted for PF10795w by SignalP 2.0

HMM (Signal peptide probability 0.152, signal anchor

probability 0.803) with cleavage site probability 0.079

between residues 40 and 41

1 probable transmembrane helix predicted for PF10795w by

TMHMM2.0 at aa 21-43"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="CAD51845.1"

/db\_xref="GI:23505063"

/db\_xref="InterPro:IPR002885"

/db\_xref="UniProtKB/TREMBL:Q81220"

/translation="MGNKMLTKKDKGKKKAVYIFLILIFDYQWSPPLVYVFCIN

KRSVLASIGKNDGPKYIRDNCKTLEIGNSIKQNSVKALANDELFLAAPDDKK

LSKWSMIGCNVLNLCGRPDINDNWKVIEKMKHNIEMQNSYLAICNYIISNNYIY

REINPHNNIHMIDNNNNLNKIDNNDYDFKQMLNKYDITLQEQDKYLHNLYL

LIDIFKHMNDNKIKLILYVYENIYQYIRKNAKELITNYKEQNISINDT

DQNLIEKDKKFTNDHEDTDSNGEQIDDMHKQNTKNNKLLKLLKDLIEQLR

NILYKWNYSMNINIKKYEETLDEBERSIYVNVNKTRELPTKLTENIKNTVC

KNCENIYFSLSKHIIIVINIIMITHQNNKELIKLHFFSILNGVDVTEGDIE

KBETESGGTNQKQGVHNVFQEQNDNNITNEMKKNKNSHNDQNGER

NHTSINENKLFDKNGYTCLLDGANIGYQNVENGRPSFLOIILKEIIRKKNNEK

PLIILPKIYHYKNSLKLHSEKANEPESSDKNEKEDILKDTHLKGRKNSHIFTN

KTIKKTPFSFEPKPTRREKDISVANGMSNNKNDTSHSEKNSYDNGHDALS

IKRIFNLDSYDVIDIIEKWEKCLIVICNVANMDDYVYILGSLKSNVNIYYIDE

LSKGFHKQTLNHNIIIDNNIMDNLLYNNREVLYVNNVYDKNSDIMLVSEENM

NREKYSIMNEQYNDKIKKNNANVLFQYKNNMKFLDKELIKYNNNIFMDTYMDM

DSGQHLVNSHEENKRI SQOEIICNNNDVLKNVNTLNDNLLIKNKSJDKDIISE

STEOQNPEKESKKNKNNNNNTFYNSIYIKCVTHMSQVQPIYITYDMKQKH

YFNEYTVILKKWKKSFISFYKQPVLSIBIEQAGNNAIDVENILQNYLPFVNLE

NFKYIDIDVYKQNVYIKINTPKYTFMSYQKDFPYMNSHNVYKYLICDFSKT"

complement(join(15222..15364,15455..15582,15715..15804,

15947..16143,16443..16553,16671..16777,16874..16928))

/gene="PF10800c"

complement(join(15222..15364,15455..15582,15715..15804,

15947..16143,16443..16553,16671..16777,16874..16928))

/gene="PF10800c"

/note="4 probable transmembrane helices predicted for

PF10800c by TMHMM2.0 at aa 7-29, 88-110, 130-152 and

162-184"

/codon\_start=1

/product="hypothetical protein"

/protein\_id="CAD51846.1"

/db\_xref="GI:23505064"

/db\_xref="UniProtKB/TREMBL:Q81219"

/translation="MFNGCINHFNILDAAFCVGIQPSFETTYGIFVSPYNSKYIKWHE

FESHACKENVALPOINYASKGFEPTMRKQLCHSINFLKNYLITLITVILIV

VKGVLLYAHFHEPPTFALKVFCNKLITLITICITITNCILLVAPVFFNSVEEFGK

THLHAGFYILLADFLTHFIIVIFLKRDKTILKVKVCELELIPWDEKRKLPHVHAFBND  
TLSPBEITNYLOGLFOQDNIENEOLEKQNFQTDVYQYVYVILNNLKNQKQMMNITPM  
" gene join(18553..20050,20090..26112)  
/gene="PFI0805w"  
CDS join(18553..20050,20090..26112)  
/gene="PFI0805w"  
/note="No pfam hit, SMART hit to SM00184, Ring finger"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD51847.1"  
/db\_xref="GI:23505065"  
/db\_xref="GOA:Q812Y8"  
/db\_xref="UniProtKB/TREMBL:Q812Y8"  
/translation="MTKIQNDIMLPFFSVSKHKNNVNTVYCRTRKRKRKRKRKKEK  
YCHRCINIRKILKYEKNLNKSNVYVYIDLKMIKNKSYENDYKFKVYKNNIKSNIN  
SNNTIKKKYKLESIPFYNDYIYSMTYIESLVITWWSNGIISLFNHNHIVGRKVT  
NHLTKNLCMNCSENIAPMDNDSLYIFNLNNIINKGLHKNVIFISNDIYEKGVKNY  
NIDYNNREIDRLSVDFTYNKLKKKKWYDMNNKKONIINVIISDEIMKYNINLNTT  
FCINPYNNNNNNNNICVLTKKTVSLIYIEYHINSVLLFKDDNINLNIWYSEFIF  
Alignment Scores:  
Pred. No.: 0.00847 Length: 341050  
Score: 120.00 Matches: 69  
Percent Similarity: 36.0% Conservative: 43  
Best Local Similarity: 22.2% Mismatches: 103  
Query Match: 7.7% Indels: 96  
DB: 13 Gaps: 19  
US-10-734-719-9 (1-291) x PFA929357 (1-341050)  
Qy 17 AspTyrSerArgLeuProAsnAspPheAspValPhe----- 28  
Db 265795 GATTATTTCGAAAAAGTAGAAGATATTCAAGTTTATTCTTTTAAACATTATATAAAAA 265736  
Qy 29 ArgCysAsnGlnPheTyr-----PheGluAspLysTyrTyrLeuGly 42  
Db 265735 GAATGTGATGAATGTGATCATCTAAATCGACCATATTATTAATAAAAAATATATCAT 265676  
Qy 43 LysLysCysLysAlaValPheTyrThrProAsnPhePheGluGlnTyrTyrThrLeu 62  
Db 265675 AACAGATCAATGATCATCATATATAAACAATTTTAT-----CGTTTAAATAAATA 265622  
Qy 63 LysHisLeuLeGlnAsnGlnGlyTyrGluThrGluLeuLeuMetCysSerAsnTyrAsn 82  
Db 265621 AAAAGT---ATAAGGATTATAAATAAATGTAATATTCTTCTGCCATTTATGGA 265565  
Qy 83 GlnAlaHisLeuGluAsnGluAsnPheVal-----LysThr 94  
Db 265564 GATACTGAAAAAGTAATTCATCTTTGTTTAAAGAAATATAGAAAAAGAAAAA 265505  
Qy 95 PheTyrAspTyrThrProAsnAlaHisLeuGlyTyrAspPhePheLysGlnLeuLysGlu 114  
Db 265504 CTAATGATTTATTAATGAT----- 265484  
Qy 115 PheAsnAlaTyrPheLysPheHisGluLeuTyrPheAsnGlnArgIleThrSerGlyVal 134  
Db 265483 -----AAATTATTGATATATAAAAAAATTGAAGAGAAATACCTCATACAAA 265439  
Qy 135 TyrMet-----CysAlaValAlaIleAlaLeuGlyTyrLysGluLeu 148  
Db 265438 TTAATGATTTCTTGATAAAGCTAATTTGTACATATAATTTAAGTGTAGGATAT---AATATT 265382  
Qy 149 TyrLeuSerGlyIleAspPheTyrClnAsnGlySerSerTyrAlaPheAspThrLysGln 168  
Db 265381 TACTTTAAGC-----GAAAAA 265364  
Qy 169 GluAsnLeuLeuLysLeuAlaProAspPheLysAsnAspArgSerHisTyrIleGlyHis 188  
Db 265363 AAAAATATATATATA-----GATATTGGAATAGTACATTTTTTTTACTTTCATTTT 265310  
Qy 189 SerLysAsnThrAspIleLysAlaLeu---GluPheLeuGluLysThrTyr----- 204

Db 265309 AATGAAATATTAATATCTTTGTTTCTTCTTATTTAAAAACAACACTTATATCTTAATT 265250  
Qy 205 -----LysIleLysLeuTyrCys---LeuCysProAsnSerLeuLeuAlaAsnPhe 220  
Db 265249 ATGTATCATCATATCTAAATATTGTGACATATGTATGAATTCATTT-----AACCTT 265196  
Qy 221 IleGluLeuAlaProAsnLeuAsnSerAsnPheIleIleGlnGluLysAsnAsnTyrThr 240  
Db 265195 TTGGAT-----AAATCATATAAAAGGTGTGTATGCATATAAAATATATGTA 265151  
Qy 241 LysAspIleLeuIleProSerSerGluAlaTyrGlyLysPheSerLysAsnIleAsnPhe 260  
Db 265150 AAG-----AAAAATGAAGATAAATATGGT-----ATAAGTAGATATATAGATCAG 265106  
Qy 261 LysLysIleLysIleLysGluAsnValTyrTyrLysLeuIleLysAspLeu---LeuArg 279  
Db 265105 AAG-----TACATGTGGATAAAATTTATGATATAAGG 265073  
Qy 280 LeuProSerAspIleLysHisTyrPheLysGly 290  
Db 265072 AATGACAAATATATAAATCATTTATTAATGGT 265040  
RESULT 147  
PFMAL13\_13/c  
WPCOMMENT  
Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509  
Fragment Name Begin End  
PFMAL13\_00 1 110000  
PFMAL13\_01 100001 210000  
PFMAL13\_02 200001 310000  
PFMAL13\_03 300001 410000  
PFMAL13\_04 400001 510000  
PFMAL13\_05 500001 610000  
PFMAL13\_06 600001 710000  
PFMAL13\_07 700001 810000  
PFMAL13\_08 800001 910000  
PFMAL13\_09 900001 1010000  
PFMAL13\_10 1000001 1110000  
PFMAL13\_11 1100001 1210000  
PFMAL13\_12 1200001 1310000  
PFMAL13\_13 1300001 1410000  
PFMAL13\_14 1400001 1510000  
PFMAL13\_15 1500001 1610000  
PFMAL13\_16 1600001 1710000  
PFMAL13\_17 1700001 1810000  
PFMAL13\_18 1800001 1910000  
PFMAL13\_19 1900001 2010000  
PFMAL13\_20 2000001 2110000  
PFMAL13\_21 2100001 2210000  
PFMAL13\_22 2200001 2310000  
PFMAL13\_23 2300001 2410000  
PFMAL13\_24 2400001 2510000  
PFMAL13\_25 2500001 2610000  
PFMAL13\_26 2600001 2710000  
PFMAL13\_27 2700001 2732359  
Continuation (14 of 28) of PFMAL13 from base 1300001 (AL844509 Plasmodium falciparum 3D

Alignment Scores:  
Pred. No.: 0.00329 Length: 110000  
Score: 119.50 Matches: 73  
Percent Similarity: 37.0% Conservative: 49  
Best Local Similarity: 22.1% Mismatches: 103  
Query Match: 7.6% Indels: 105  
DB: 12 Gaps: 19

US-10-734-719-9 (1-291) x PFMAL13\_13 (1-110000)

Qy 23 AsnAspPheAspValPheArgCysAsnGlnPheTyrPheGluAspLysTyr----- 39  
Db 62254 GATGACCATGATGATGATCGTACAAACCGCATTTTACAATGGGATAAATTTCTTCCCAT 62195  
Qy 40 -----TyrLeuGlyLysCysLysAlaVal--- 48







